

# NMR-spektroskopische Charakterisierung der nicht stabil faltenden Villin-Headpiece36-Mutante mut8

Dissertation

zur Erlangung des Doktorgrades

Dr. rer. nat.

der Fakultät Biologie und Geographie

der Universität Duisburg-Essen

Campus Essen

vorgelegt von

Anja Matena

aus Essen

April 2010

Die der vorliegenden Arbeit zu Grunde liegenden Experimente wurden in der Abteilung Strukturelle und Medizinische Biochemie am Zentrum für Medizinische Biotechnologie der Universität Duisburg-Essen durchgeführt.

1. Gutachter: Prof. Dr. Peter Bayer
2. Gutachter: Prof. Dr. Daniel Hoffmann
3. Gutachter: Prof. Dr. Raphael Stoll

Vorsitzender des Prüfungsausschusses: Prof. Dr. Hynek Burda

Tag der mündlichen Prüfung: 27. Juli 2010

# Inhaltsverzeichnis

<b>Abkürzungsverzeichnis</b>	<b>6</b>
<b>Abbildungsverzeichnis</b>	<b>8</b>
<b>Tabellenverzeichnis</b>	<b>11</b>
<b>1. Einleitung</b>	<b>12</b>
1.1. Villin Headpiece 36 (HP36) als Modell in der Bioinformatik . . . . .	12
1.2. Der evolutionäre Algorithmus EMOS und HP36 . . . . .	16
1.3. EMOS als neues Zusatzwerkzeug für Peptid-Design in der medizinischen Forschung . . . . .	20
1.3.1. Ein Beispiel für die Anwendung von EMOS . . . . .	20
1.4. Zielsetzung . . . . .	24
<b>2. Materialien und Methoden</b>	<b>26</b>
2.1. Materialien . . . . .	26
2.1.1. Chemikalien . . . . .	26
2.1.2. DNA- und Protein-Marker . . . . .	27
2.1.3. Puffer und Lösungen . . . . .	27
2.1.4. Nährmedien und Antibiotika . . . . .	29
2.1.5. Bakterienstämme . . . . .	30
2.1.6. Plasmid . . . . .	30
2.1.7. Enzyme . . . . .	31
2.1.8. Verbrauchsmaterialien . . . . .	31
2.1.9. Geräte . . . . .	31
2.1.10. Software . . . . .	32
2.2. Peptidsynthese . . . . .	33
2.3. Mikro- und Molekularbiologische Methoden . . . . .	33
2.3.1. Klonierung von mut8 . . . . .	33
2.3.2. Überprüfung der korrekten Klonierung . . . . .	35
2.3.3. Expression von mut8 in <i>E. coli</i> Rosetta (DE3) . . . . .	36
2.4. Biochemische Methoden . . . . .	36
2.4.1. Zellaufschluss . . . . .	36
2.4.2. Proteinreinigung . . . . .	36
2.4.3. Proteinkonzentrationsbestimmung . . . . .	38
2.4.4. SDS-PAGE . . . . .	38
2.5. NMR-Spektroskopie . . . . .	39
2.5.1. Probenvorbereitung . . . . .	39

2.5.2.	Gefriertrocknung von mut8 . . . . .	39
2.5.3.	Messbedingungen . . . . .	39
2.5.4.	Prozessierung der Spektren . . . . .	40
2.5.5.	Auswertung der Spektren . . . . .	41
2.6.	Strukturaufklärung und Dynamik . . . . .	44
2.6.1.	Das Xplor-NIH-Kraftfeld . . . . .	44
2.6.2.	Erstellen der Koordinatendateien . . . . .	44
2.6.3.	<i>Simulated Annealing</i> -Läufe . . . . .	45
2.6.4.	Verfeinerung der Struktur . . . . .	45
2.6.5.	Analyse der Ensembles . . . . .	45
2.6.6.	Dynamik-Berechnungen . . . . .	46
<b>3.</b>	<b>Ergebnisse</b>	<b>47</b>
3.1.	NMR-Spektroskopie von mut8 . . . . .	48
3.1.1.	Spektrenaufnahme und -zuordnung . . . . .	48
3.1.2.	Sekundärstrukturanalyse . . . . .	58
3.1.3.	Tertiärstrukturberechnung . . . . .	63
3.2.	Klonierung von mut8 . . . . .	68
3.3.	Expression und Reinigung von mut8 . . . . .	70
3.3.1.	Expression und GSH-Affinitätschromatographie . . . . .	70
3.3.2.	Optimierung der Expression und Reinigung . . . . .	71
3.4.	Dynamik von mut8 . . . . .	73
3.4.1.	Aufnahme des SOFAST-HMQC und Zuordnung . . . . .	73
3.4.2.	Aufnahme des HSQCNOE und Berechnung des hetNOE . . . . .	74
<b>4.</b>	<b>Diskussion</b>	<b>78</b>
4.1.	mut8 hat eine gut definierte Sekundärstruktur . . . . .	79
4.1.1.	NOEs und CSI – Resultate aus der NMR-Spektroskopie . . . . .	79
4.1.2.	NOEs und DSSP – Resultate aus der Bioinformatik . . . . .	80
4.1.3.	Die Diederwinkel von mut8 und new2 . . . . .	85
4.2.	Die Tertiärstruktur von mut8 . . . . .	87
4.3.	Expression von mut8 und heteronukleare NMR-Spektroskopie . . . . .	93
4.4.	Bei mut8 handelt es sich um ein dynamisches Peptid . . . . .	96
<b>5.</b>	<b>Ausblick</b>	<b>103</b>
<b>6.</b>	<b>Zusammenfassung</b>	<b>106</b>
	<b>Literaturverzeichnis</b>	<b>108</b>
<b>A.</b>	<b>Protokolle</b>	<b>116</b>
<b>B.</b>	<b>Ergebnisse</b>	<b>125</b>
	<b>Danksagung</b>	<b>141</b>



<b>Lebenslauf</b>	<b>143</b>
<b>Erklärungen</b>	<b>144</b>

# Abkürzungsverzeichnis

1D	eindimensional
2D	zweidimensional
Å	Ångström
Abb.	Abbildung
AS	Aminosäure
bidest	zweifach destilliertes Wasser
BSA	Rinderserumalbumin
bzgl.	bezüglich
bzw.	beziehungsweise
ca.	circa
C-Terminus	Carboxy-Terminus
COSY	<i>correlated spectroscopy</i>
CSI	<i>chemical shift index</i>
d. h.	das heißt
evtl.	eventuell
DSS	2,2-Dimethyl-2-silapentan-5-sulfonat
DTT	Dithiotreitol
<i>E. coli</i>	<i>Escherichia coli</i>
FID	<i>free induction decay</i>
FZ	Fourier-Transformation
ggü.	gegenüber
Gl.	Gleichung
HET	<i>heteronuclear</i>
HMQC	<i>heteronuclear multiple quantum coherence</i>
HP36	Villin Headpiece 36
HSQC	<i>heteronuclear single quantum coherence</i>
K	Kelvin
Kan	Kanamycin
Kap.	Kapitel
kDa	kilo-Dalton
Lsg.	Lösung
M	Molar (mol/l)
mAU	milli Absorptionsunits
MD	Molekulardynamik
MHz	Megahertz
min	Minute
min.	mindestens
mut8	Achtfachmutante des HP36

NMR	<i>nuclear magnetic resonance</i>
NOE	<i>nuclear Overhauser effect</i>
	bzw. Kreuzresonanzen im NOESY-Spektrum
NOESY	<i>nuclear Overhauser enhancement spectroscopy</i>
ns	Nanosekunde
N-Terminus	Amino-Terminus
OD	optische Dichte
PAGE	Polyacrylamidgelelektrophorese
ps	Picosekunde
ppm	<i>parts per million</i>
RMSD	<i>root mean square deviation</i>
RMSF	<i>root mean square fluctuation</i>
rpm	<i>rounds per minute</i>
RT	Raumtemperatur
SA	<i>simulated annealing</i>
SDS	Natriumdodecylsulfat
s. o.	siehe oben
Tab.	Tabelle
TOCSY	<i>total correlation spectroscopy</i>
TPPI	<i>time proportional phase incrementation</i>
u. a.	unter anderem
ü. N.	über Nacht
vdW	van der Waals-Kräfte
vgl.	vergleiche
w/v	Gewicht pro Volumen
w/w	Gewicht pro Gewicht
z.B.	zum Beispiel
$\delta$	chemische Verschiebung
$\eta$	hetNOE
$\mu$	mikro
$\phi$	Phi-Diederwinkel (C-N-C $_{\alpha}$ -C) (Peptidrückgrat)
$\psi$	Psi-Diederwinkel (N-C $_{\alpha}$ -C-N) (Peptidrückgrat)

Aus dem Englischen übernommene Fachausdrücke sind kursiv gesetzt. Die Aminosäurenomenklatur wurde nach den Vorschlägen der IUPAC-IUB-Kommision für Biochemische Nomenklatur übernommen (Eur. J. Biochem., 1984, 138: 9-37).

# Abbildungsverzeichnis

1.1.	Villin Headpiece 36 aus <i>Gallus gallus</i> (pdb-ID: 1vii). . . . .	13
1.2.	A) Faltungs-Energielandschaft von HP35 bei 300 K. (B) Schnappschüsse der Faltungssimulation der K65NLE/N68H/K70NLE-Mutante von HP35 nach 0, 1, 10, 70, 100 und 700 ns. . . . .	14
1.3.	Strukturen von Villin Headpiece (pdb-ID: 1unc) und Advillin Headpiece (pdb-ID: 1und) aus <i>Homo sapiens</i> . . . . .	15
1.4.	Schematische Darstellung des evolutionären Algorithmus EMOS. . . . .	18
1.5.	Vergleich zwischen Modellstruktur (grün) und experimentell bestimmter Struktur der HP36-Mutante G34L (1ppz). . . . .	19
1.6.	Sequenz-Alignment von stabilen Individuen. . . . .	19
1.7.	Schematische Darstellung des HIV-Infektionsmechanismus. . . . .	21
1.8.	Strukturelle Darstellung der Bindung des N-Terminus von CCR5 an gp120. . . . .	22
2.1.	Verwendete DNA- und Protein-Marker. . . . .	27
2.2.	Schematische Darstellung des pET41b_mod-Vektors (links) mit einklonierter PreScission-Protease-Erkennungssequenz (rechts). . . . .	31
2.3.	DNA-Konstrukt zur Expression von mut8. . . . .	35
2.4.	Typische Signalmuster der einzelnen Aminosäuren. . . . .	41
2.5.	Prinzip des <i>Chaintracing</i> -Verfahrens nach Wüthrich, 1986. . . . .	42
3.1.	1D- <sup>1</sup> H-Spektrum des synthetisch hergestellten mut8. . . . .	49
3.2.	Schematische Darstellung eines Tripeptids aus zwei Phenylalaninen und einem Leucin und die jeweiligen Kopplungen, die sich im COSY-, TOCSY- und NOESY-Spektrum ergeben würden. . . . .	50
3.3.	Die drei übereinandergelegten 2D-Spektren und Ausschnitte aus der <i>Fingerprint</i> -Region des COSYs (oben rechts) sowie dem Bereich der H <sub>N</sub> -H <sub>N</sub> -Kopplungen (unten rechts). . . . .	51
3.4.	Überlagerung von TOCSY- und NOESY-Spektrum des synthetischen mut8-Peptids. . . . .	52
3.5.	Darstellung von Prolin innerhalb einer Peptidkette aus Leu-Pro-Leu. . . . .	53
3.6.	Die drei übereinandergelegten 2D-Spektren (vgl. 3.3) und ein Ausschnitt aus der Aliphaten-Region mit einer Darstellung von Prolin und dessen Bezeichnungen der Protonen nach IUPAC-Nomenklatur (rechts). . . . .	54
3.7.	ollständige Zuordnung von L35. . . . .	55

3.8.	Die drei übereinandergelegten 2D-Spektren (vgl. Abb. 3.3) und ein Ausschnitt (oben), welcher den ersten Zuordnungsschritt der Aminoprotonen von Q26 zeigt. Unten ist der oben rechts markierte Bereich vergrößert dargestellt, links davon befindet sich das Glutamin mit Bezeichnung der Protonen nach IUPAC-Nomenklatur. . . . .	56
3.9.	Die drei übereinandergelegten 2D-Spektren (vgl. Abb. 3.3) und ein Ausschnitt (oben), welcher den ersten Zuordnungsschritt der Ringprotonen von W34 zeigt. . . . .	57
3.10.	Darstellung der Diederwinkel $\phi$ und $\psi$ innerhalb einer Peptidkette. . . .	58
3.11.	Darstellung der Konnektivitäten innerhalb von mut8. . . . .	61
3.12.	Darstellung des chemischen Verschiebungsindex (CSI) für mut8. . . .	62
3.13.	Anzahl und Art der zur Tertiärstrukturberechnung eingesetzten NOE-Beschränkungen pro Aminosäure. . . . .	64
3.14.	Energieprofil der erhaltenen Strukturenscharen nach Verwendung des <i>sa.inp</i> (schwarz)- und des <i>ref.inp</i> (rot)-Protokolls. . . . .	65
3.15.	Darstellung der für mut8 berechneten Strukturen des Ensembles. . . . .	67
3.16.	Ramachandranplot des Ensembles für mut8 (RamPage [1]). . . . .	67
3.17.	1,5 %iges Agarose-Gel zur Verifizierung der erfolgreichen Klonierung von mut nach Testrestriktion des Vektors mit ApaI und XhoI. . . . .	69
3.18.	DNA-Sequenz von mut8 (blau unterlegt) nach der Sequenzierung (GATC) im Vergleich zur elektronischen Kontroll-Klonierung. . . . .	69
3.19.	Elutionsverlauf der GSH-Reinigung mit dem dazu passenden 12,5 %igen SDS-Gel nach einer Coomassie-Färbung. . . . .	70
3.20.	12,5 %ige SDS-Gele nach GSH-Aufreinigung nach herkömmlicher Expression und optimierter Kurzexpression. . . . .	71
3.21.	Gelfiltrationsprofil nach Trennung der GST und mut8 und das dazugehörige 15 %ige SDS-Gel. . . . .	72
3.22.	SOFAST-HMQC mit Zuordnung. . . . .	74
3.23.	Abhängigkeit der maximalen theoretischen NOE-Intensität von $\omega_x \cdot \tau_c$ für X( <sup>1</sup> H)-Experimente; X= <sup>1</sup> H (homonuklearer NOE), <sup>13</sup> C und <sup>15</sup> N. . .	76
3.24.	Darstellung der Berechnung von $I_{sat}/I_{eq}$ . . . . .	77
4.1.	Schematische Zeichnung einer Trajektorie in Vektor-Darstellung (links) und Übertragung davon auf die zwei möglichen Trajektorien eines Peptids in einem Raum (gefüllt mit Wasser, bei Raumtemperatur) ausgehend von einer Startstruktur (grau). . . . .	81
4.2.	Ein Beispiel der Distanz-Bestimmungen über NOEs. . . . .	82
4.3.	DSSP-Berechnungen für je 1000 Strukturen zweier Trajektorien von new2 (links, Mitte) und des Ensembles von mut8 (rechts). . . . .	84
4.4.	Darstellung der Diederwinkelbestimmung von Aspartat 6 in Trajektorie 1 und 2. . . . .	85
4.5.	Darstellung der new2/mut8-Sequenz für Trajektorie 1 und 2. . . . .	86
4.6.	Darstellung der für mut8 berechneten Strukturen des Ensembles im Vergleich zu den Strukturen für r1g11i, welche mit den bioinformatisch ermittelten $\phi$ - und $\psi$ -Winkeln berechnet wurden. . . . .	87

---

4.7.	Mit Hilfe des Programms PyMol überlagerte Headpiece-Strukturen und ein Alignment der Primärsequenz. . . . .	88
4.8.	Alignment von 1vii und r1g11i1, erstellt mit ClustalW [2], dargestellt und bearbeitet mit Jalview [3]. Blau unterlegt sind alle Aminosäuren, welche in beiden Sequenzen übereinstimmen. . . . .	88
4.9.	Darstellung einer mut8-Struktur des Ensembles im <i>cartoon</i> -Modell mit Hilfe von PyMol. . . . .	89
4.10.	Schematische Darstellung von dipolaren Kopplungen zwischen Protonen der Aminosäuren 5 und 6 (AS 5, 6) und Aminosäuren 21 bis 23 (AS 21, 22, 23) innerhalb einer Peptidkette. . . . .	90
4.11.	Darstellung von 1vii. Hervorgehoben sind alle für den hydrophoben Kern wichtigen Aminosäuren sowie die beiden in r1g11i1 mutierten Residuen F11D und F18K in rot. . . . .	91
4.12.	Darstellung von Villin Headpiece 35 (pdb-ID: 1unc) mit dem PxWK-Motiv (links). . . . .	92
4.13.	Die Amidregion der NOESY-Spektren von HP36 (1unc) (links), von dessen HP36-P22A-Mutante (rechts) und von mut8. . . . .	94
4.14.	Darstellung der Zeitskala für molekulare Dynamik (blau) und NMR-spektroskopische Methoden (grün) (verändert nach [4]). . . . .	96
4.15.	Ausschnitt des SOFAST-HMQC und des HSQCNOE <sub>eq</sub> . . . . .	97
4.16.	Darstellung der <sup>1</sup> H- <sup>1</sup> H Spin-Diffusion und von Amid-Wasser-Protonen Austauschwirkungen in einer ausgestreckten Peptidkette und einem globulären, strukturierten Protein. . . . .	99
4.17.	Darstellung der RMSF-Werte für den HP36-Wildtyp 1vii (schwarz) und new2 aus 2 Trajektorien (rot, blau). . . . .	100
4.18.	RMSD-Werte über 10 ns MD-Simulation von drei Peptiden. . . . .	101
5.1.	Analysemethode für Sequenzen nach einem neuen EMOS-Durchlauf. . .	104

# Tabellenverzeichnis

2.3.	Programm der Chromatographie von mut8, die verwendete GSH-Säule hat ein Volumen von 1 ml. Zwei dieser Säulen wurden direkt hintereinander geschaltet, somit ergibt sich ein Gesamt-Säulenvolumen von 2 ml. . . . .	37
2.4.	Programm der Gelfiltrationschromatographie von mut8, die verwendete Säule HiLoad 16/60 Superdex 75pg (GE-Healthcare) hat ein Volumen von 125 ml. . . . .	37
2.5.	Zusammensetzung unterschiedlich konzentrierter Polyacrylamidgele. . .	38
2.6.	NMR-Experimente und Parameter. 1: Zahl der gemessenen <i>scans</i> , 2: spektrale Breite in Hz, 3: Zahl der realen Punkte in der jeweiligen Dimension. . . . .	40
3.1.	Auflistung aller zugeordneten Protonen mit Hilfe der 2D-Spektren. Die Bezeichnung der Atome erfolgt nach der IUPAC-Nomenklatur. Mit eingefügt sind vorausgreifend die Zuordnungen der Stickstoff-Resonanzen (siehe Kap. 3.4.1). . . . .	59
3.2.	Strukturelle Statistik für das energieminierte Ensemble aus 10 Strukturen berechnet für mut8 mit Xplor-NIH. . . . .	66
4.1.	Auflistung der für new2/mut8 aufgrund der NOEs, des CSI und der DSSP ermittelten $\alpha$ -helikalen Bereiche. . . . .	84

# 1. Einleitung

Einfache Lebensformen bis hin zu komplexen Organismen werden durch evolutionäre Prozesse über Jahrmillionen hinweg an ihre Umwelt- und Lebensbedingungen angepasst. Die Evolution ist somit ein natürliches Vorbild für ein eigentlich schwieriges Optimierungsproblem. Die Mutation und Selektion bilden dabei zwei biologische Prinzipien, auf die sich die Vorgehensweise dieses natürlichen Optimierungsverfahrens zurückführen lässt [5].

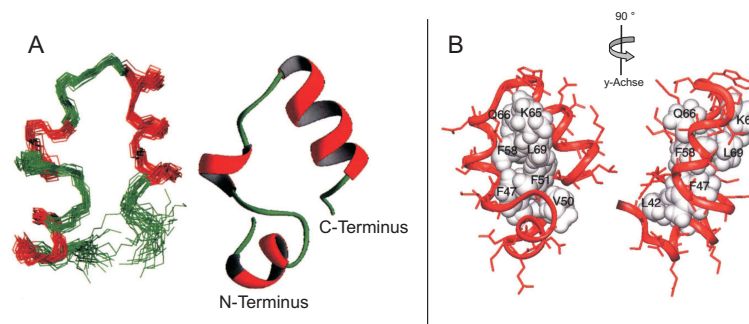
Das Prinzip der Evolution diente schon in den 1960er Jahren Forschergruppen als Vorlage für die spätere Entwicklung von evolutionären Algorithmen (EA) [6, 7]. EAs simulieren evolutionäre Prozesse und erschaffen zunächst eine Menge von Individuen (= Lösungen). Darauf folgt die Erschaffung von Nachkommen durch Mutation und sowohl die Startpopulation als auch die Nachkommen werden durch eine Fitnessfunktion bewertet. Nach der Bewertung der Individuen erfolgt die Selektion, bei der schlechte Lösungsvorschläge aussortiert und gute beibehalten werden. Die Resultate der Selektion stellen dann die hinsichtlich bestimmter Kriterien optimierten Individuen dar. Diese Schritte der Mutation, Bewertung und Selektion werden so lange wiederholt, bis ein vorher definiertes Abbruchkriterium erfüllt ist. Dies kann eine definierte Anzahl an Durchläufen sein oder ein Stillstand des Wertes der besten Lösung [8]. Bei der Mutation handelt es sich um einen ungerichteten Prozess, wodurch Alternativen und Varianten erzeugt werden. Die Selektion ist für die Steuerung der Evolution in eine bestimmte Richtung zuständig.

Evolutionäre Algorithmen können dazu verwendet werden, Proteine oder Peptide *in silico* auf bestimmte Kriterien hin zu optimieren. So kann z. B. deren strukturelle Stabilität verbessert oder auf eine zugängliche Tasche für hochaffine Bindung eines Liganden hin optimiert werden [9]. Um solche Algorithmen testen und anwenden zu können, werden geeignete Modellsysteme benötigt. Dafür kommen vor allem kleine Proteine und Peptide in Frage, welche *in vitro* bereits gut untersucht sind. Eines dieser Peptide ist Villin Headpiece 36 (HP36).

## 1.1. Villin Headpiece 36 (HP36) als Modell in der Bioinformatik

Villin, ein bereits 1979 mit einem Molekulargewicht von 95 kDa beschriebenes Protein [10], besteht aus einem Kern- (ca. 87 kDa) und einem C-terminal gelegenen Headpiece-Bereich (ca. 8,5 kDa), welcher auch als HP76 bezeichnet wird, da er aus 76 Aminosäuren besteht [11, 12]. Villin ist gewebespezifisch und liegt in Assoziation mit dem mikrovillären Aktin des Darm- und Nieren-Bürstensaums vor. *In vitro* konnte



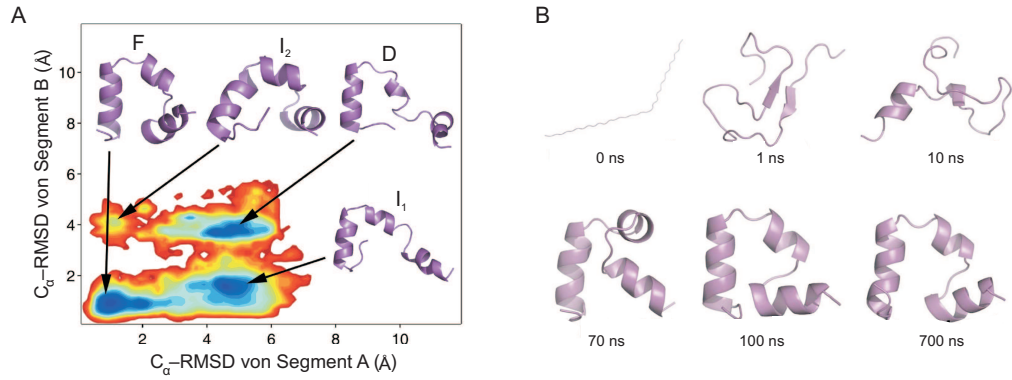


**Abbildung 1.1.:** Villin Headpiece 36 aus *Gallus gallus* (pdb-ID: 1vii). (A) Überlagerung aller 29 akzeptierten NMR-Strukturen (links) und *ribbon*-Darstellung der minimierten und gemittelten NMR-Struktur von HP36. (B) Hydrophobe Kern-Reste von HP36. Alle als Kugeln dargestellten Aminosäuren sind zu >70 % nicht Lösungsmittel-zugänglich. Das Rückgrat ist als rote *ribbon*-Darstellung gezeigt, weitere Seitenketten als *sticks*. Verändert nach [17].

gezeigt werden, dass Villin F-Aktin bündelt und auflöst [13]. Zusätzlich besitzt Villin drei Phosphatidylinositol-4,5-bisphosphat (PIP2) Binderegionen, zwei davon sind im Kernbereich lokalisiert, eine am C-terminalen Bereich [14]. Die Kerndomäne löst in einer  $\text{Ca}^{2+}$ -abhängigen Reaktion Aktinfilamente auf, während die Aktinfilament-Bündelung  $\text{Ca}^{2+}$ -unabhängig abläuft und abhängig von der F-Aktin-Bindekapazität des C-terminalen Bereiches ist [15]. HP76 besitzt C-terminal eine Domäne, die als HP35 bekannt ist und eine unabhängig stabil faltende Untereinheit aus 35 Aminosäuren darstellt [16]. Die Struktur dieser stabil faltenden, kleinen Untereinheit HP36 – die Zahl 36 ergibt sich dadurch, dass die rekombinante Form von HP35 noch ein Methionin am N-Terminus trägt – wurde 1997 von McKnight *et al.* mittels NMR gelöst [17].

HP36 besteht aus drei  $\alpha$ -Helices, welche über Schleifen miteinander verbunden sind (siehe Abb. 1.1). Der hydrophobe Kern des Peptids wird von Aminosäuren aller drei Helices gebildet. Die Aminosäurereste L42, F47, V50, F51, A57, F58, K65, Q66, L69 und G74 (Nummerierung nach HP76) sind über 70 % nicht Lösungsmittel-zugänglich, F47 ist dabei die verdeckteste Aminosäure und kann sogar zu 95 % nicht vom Lösungsmittel erreicht werden. Die Aminogruppen der Seitenketten von K65 und Q66 sind von den genannten Aminosäuren sehr Solvens-exponiert. In homologen Headpiece-Sequenzen aus dem Huhn [12], dem Menschen [18] oder der Maus (Genbank *accession number* M98454) und sogar dem proto-Villin aus *Dictyostelium* [19] sind die hydrophoben Kern-Residuen F47, F58, K65 und L69 konserviert. Die ebenfalls konservierten L42 und M53 bilden Kontakte mit dem hydrophoben Kern und schirmen diesen vom Solvens ab, während das konservierte W64 sich genau über dem N-Terminus an der C-terminalen Helix befindet und Kontakte mit dem L61 und P62 bildet, sodass auch K65 vom Lösungsmittel abgeschirmt wird. L75 und F76 sind ebenfalls konserviert in den Headpiece-Sequenzen.

Bereits 1998 geriet HP36 in das Interesse der Bioinformatik. So führten Duan *et al.*, 1998 eine Molekulardynamik-Simulation durch, in der die Faltung von HP36 ausge-

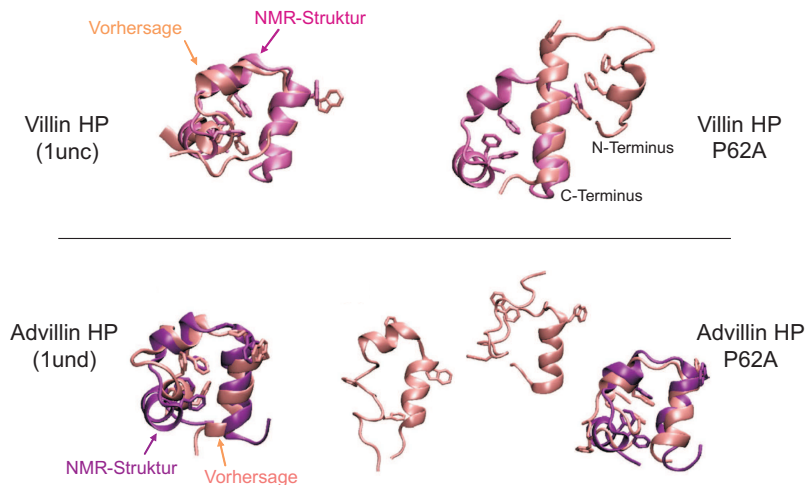


**Abbildung 1.2.:** (A) Faltungs-Energielandschaft von HP35 bei 300 K. Segment A umfasst zusammengelagerte Helix I und II, Segment B umfasst Helix II und III. Über der Energielandschaft sind repräsentativ Strukturen für die 4 Faltungsstadien abgebildet. F: gefalteter Zustand, D: denaturierter Zustand,  $I_1$ : größerer intermediärer Zustand,  $I_2$ : kleinerer intermediärer Zustand. Verändert nach [21]. (B) Schnappschüsse der Faltungssimulation der K65NLE/N68H/K70NLE-Mutante von HP35 nach 0, 1, 10, 70, 100 und 700 ns. Verändert nach [22].

hend von einer ausgestreckten Peptidkette zu einer Tertiärstruktur simuliert wurde. Für die Simulation wurden als Startkoordinaten die NMR-Struktur von HP36 (1vii) verwendet und HP36 in einer 1,0 ns-Simulation in Wasser bei 1000 K denaturiert. Nach 20 ns der Simulation bei 1000 K bildete sich eine semistabile, kompakte Struktur. Im Anschluss daran erfolgte eine Äquilibrierungsphase von 1 ns bei 200 K, bevor dann die Simulation zur Faltung von HP36 über 200 ns bei 300 K durchgeführt wurde [20].

Lei *et al.* haben 2007 die Faltungslandschaft von HP35 untersucht [21]. MD-Simulationen liefen bei 300 K über 200 ns und HP35 faltete einheitlich zum nativen Zustand. Der niedrigste  $C_\alpha$ -RMSD-Wert (*Root Mean Square Deviation*) zur Röntgenkristallographie-Struktur (pdb-ID: 1yrf) betrug nur 0,46 Å, die gemittelte niedrigste  $C_\alpha$ -RMSD der größten HP35-Population betrug 1,78 Å. Dieses Ergebnis wurde bis dahin nicht in *ab initio*-Simulationen erzielt. Die Faltungslandschaft von HP36 konnte schließlich in vier verschiedene Faltungs-Regionen eingeteilt werden: In die native (F), die denaturierte (D) und zwei Intermediat-Regionen ( $I_1$ ,  $I_2$ ). Nativer Zustand und der größere intermediäre Zustand  $I_1$  sind lediglich durch eine kleine Energiebarriere voneinander getrennt, wohingegen eine größere energetische Barriere zwischen dem Intermediat  $I_1$  und denaturierten Zustand besteht.

Ein Jahr später wurde von derselben Arbeitsgruppe der Faltungsmechanismus der K65NLE/N68H/K70NLE-Mutante (statt K65 wurden die drei Aminosäuren NLE eingefügt, analog dazu wurde K70 durch NLE ersetzt) simuliert und diese mit einer Zeitkonstante von 0,6  $\mu$ s als das schnellste sich faltende Protein deklariert [22]. Die Mutante faltet in der Zeit zu einer nativen Struktur mit einer  $C_\alpha$ -RMSD von 0,5 Å. Gleichzeitig ist diese Mutante eher ein *Downhill-folder* als der Wildtyp 1vii, d. h. sie begegnet keiner signifikanten Energiebarriere im Faltungsprozess, wie es bei HP35 der



**Abbildung 1.3.:** Strukturen von Villin Headpiece (pdb-ID: 1unc) und Advillin Headpiece (pdb-ID: 1und) aus *Homo sapiens*. Oben links ist in violett die NMR-Struktur von 1unc dargestellt und in rosafarben die Vorhersage, oben rechts die vorhergesagte Struktur für die P62A-Mutante von 1unc (rosafarben) in Überlagerung mit der NMR-Struktur (violett). Deutlich zu erkennen ist die Abweichung vor allem im N-terminalen Bereich der Mutante zu 1unc. Unten links ist in violett die NMR-Struktur von 1und gezeigt und in rosafarben die Vorhersage. Unten rechts sind die vorhergesagten Strukturen für die P62A-Mutante von 1und aus drei berechneten Clustern dargestellt. Der größte Cluster (ganz rechts) zeigt sehr große Ähnlichkeit zu 1und selbst.

Fall ist. Weitere Untersuchungen der Sekundärstruktur zeigten, dass die mutierten Reste höhere intrinsische Präferenz zur Bildung einer  $\alpha$ -Helix haben, was die Bildung von Helix III und dem Intermediat, welches die gefalteten Bereiche Helix II/III besitzt, erleichtert. Abbildung 1.2 zeigt sowohl die Energielandschaft des HP35-Wildtyps (A) als auch den Faltungsablauf der K65NLE/N68H/K70NLE-Mutante (B).

Neben diesen und zahlreichen anderen Faltungsstudien zu HP36 wurde von Piana *et al.* (2008) der Effekt einer Punktmutation auf die Faltung von HP36, Advillin-Headpiece und deren P62A-Mutanten vorhergesagt [23]. Über eine *ab initio*-Simulation wurde gezeigt, dass die Punktmutation P62A auf Villin Headpiece (1unc) dramatische Auswirkungen hat, während dieselbe Mutation auf die Faltung von Advillin Headpiece (pdb-ID: 1und) nur wenig Einfluss besitzt. 1unc und 1und besitzen einen Sequenzunterschied von 58 %, haben jedoch eine nahezu identische Tertiärstruktur (Rückgrat-RMSD: 0,25 Å) [24] und ähnliche thermale Stabilität (Entfaltungs-Temperatur: 321 K (1unc) und 336 K (1und)) [25]. Alle Residuen, welche den hydrophoben Kern bilden und für die Stabilität der Faltung verantwortlich sind, sind sowohl in 1unc als auch in 1und konserviert [26]. Abbildung 1.3 zeigt die Strukturen von Villin Headpiece (1unc), Advillin Headpiece (1und) als NMR-Struktur und Vorhersage sowie deren vorhergesagte P62A-Strukturen.

## 1.2. Der evolutionäre Algorithmus EMOS und HP36

Während alle bisher erwähnten bioinformatischen Arbeiten zumeist darauf abzielten, die Faltung von HP36 und Mutanten zu untersuchen und vorherzusagen, wurde 2005 von Hohm *et al.* der evolutionäre Algorithmus EMOS (*Evolutionary Multiobjective Optimization System*) vorgestellt, welcher *in silico*-Peptiddesign ermöglichen sollte [27]. Und zwar sollte damit gezielt eine Verbesserung von Peptiden oder kleinen Proteinen erreicht werden, um diese strukturell stabiler zu machen und hinsichtlich ihrer biologischen Funktion zu optimieren. Dadurch könnten in der medizinischen Forschung Medikamente auf Peptid-Basis schnell und ohne aufwändige *screening*-Verfahren optimiert werden. Auf den medizinischen Hintergrund wird im nachfolgenden Kapitel (siehe Kap. 1.3) näher eingegangen.

Eine vorher durch eine F18K-Mutation strukturell destabilisierte Villin Headpiece 36-Mutante sollte über den evolutionären Algorithmus EMOS stabilisiert werden, d. h. dieselbe oder eine optimierte stabile Tertiärstruktur sollte wieder hergestellt werden. Ein solcher Versuch und dessen darauf folgende Versuchsreihen gehören in die Kategorie *in silico*-Proteinstruktur-Vorhersage/Proteinfaltung/*Drug-Design*, wobei gerade in den letzten Jahren ein großer Fortschritt und ein anwachsendes Interesse an dieser Art von Methode verzeichnet werden kann [28, 29, 30, 31].

Hohm *et al.*, 2005 wählten die F18K-Mutante von HP36, dessen Struktur bereits bekannt war (siehe Kap. 1.1, [16]). F18 entspricht hierbei dem F58 aus HP36 (1vii). Diese Mutante war nachweislich instabil, konnte also ihre native Faltung nicht mehr aufrecht erhalten [26]. Der Algorithmus war nicht nur in der Lage, die Mutante wieder auf ihren ursprünglichen Zustand hin zu restabilisieren, sondern fügte noch weitere Mutationen in die HP36-Sequenz ein, die dieses Peptid noch stabiler machten. Eine der rückmutierten und zusätzlich stabilisierten Peptide war die G34L-Mutante von HP36.

Der evolutionäre Algorithmus EMOS startet mit einer Anzahl an Kopien einer Sequenz mit bekannter Struktur (Individuen) und führt schrittweise einen Optimierungsprozess durch, wodurch in jedem Schritt die Individuen leicht verändert werden. Die Veränderungen bestehen dabei aus einzelnen Punktmutationen. Der Mutationsort und die Mutationsart werden über eine Methode festgelegt, die sich *in vitro alanine scan* nennt. Bei dieser Methode werden Reste in einem Peptid oder Protein identifiziert, die eine bestimmte biologische Funktion besitzen [32, 33]. Jede identifizierte Aminosäure wird dann durch ein Alanin ausgetauscht. Alanin ist mit der einzelnen Methylgruppe als Seitenkette die kleinste und neutralste Aminosäure (ausgenommen Glycin), somit bedeutet der Austausch einer Aminosäure durch Alanin, dass die ausgetauschte Seitenkette auf ein Minimum reduziert wird. Kann die jeweilige Aminosäure nun durch ein Alanin ersetzt werden, ohne dass sich dieser Austausch negativ auf das Peptid/Protein auswirkt, z. B. zu einer Destabilisierung führt, dann besteht die Möglichkeit, dass diese Aminosäure bzgl. des betrachteten Aspekts nicht essentiell ist.

Peptide und Proteine besitzen meist einige „Schlüsselresiduen“, welche wichtig sind für deren biologische Funktion an Oberflächen exponierten Schleifen. Daher erlaubt EMOS einen Benutzer-definierten Ausschluss bestimmter Aminosäuren aus Schleifenregionen aus dem Mutationsprozess.

Der Algorithmus wird zunächst dahingehend getestet, zwei verschiedene Attribute eines Peptids zu optimieren: Zum einen sollte die strukturelle Stabilität verbessert und zum anderen die Erreichbarkeit einer Binderegion eines kleinen Liganden optimiert werden. Es ist relativ schwierig, in einer allgemeinen Simulation die Erreichbarkeit einer Bindetasche für den Bindungspartner zu berechnen. Daher wird bei EMOS eine einfache Annäherung verwendet, die die Erreichbarkeit der Binderegion statt zu einem bestimmten Liganden zu Wasser definiert, da eine Binderegion meistens relativ Solvens-exponiert vorliegt.

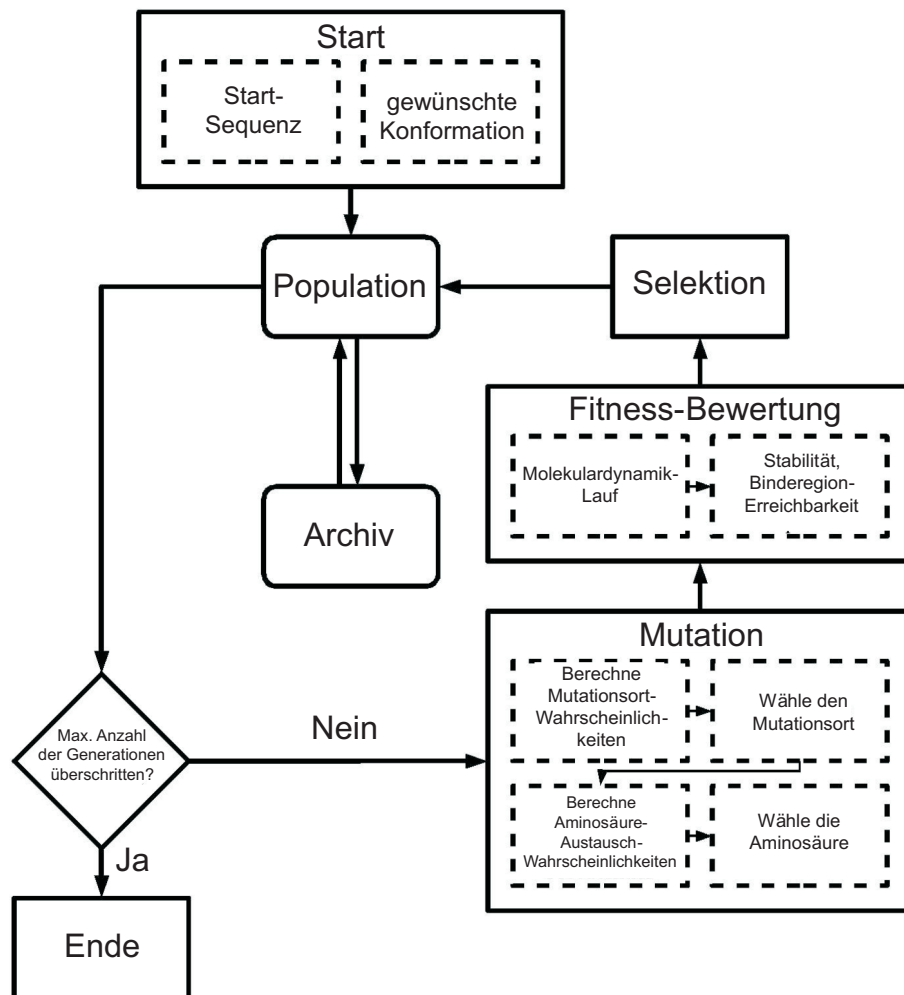
Nachdem dann die Nachkommen der anfangs genannten Individuen bewertet wurden, findet ein Selektionsprozess statt. Die folgende Generation, mit der letztendlich weiter gearbeitet wird, wird aus einem Satz der Nachkommen und der alten Generation über einen „Wettkampf“ [34] oder über Elitismus [35] ausgewählt. Der Gewinner eines solchen „Wettkampfes“ wird nach der Dominanz-Selektion bestimmt, wobei die Parameter Stabilität und Erreichbarkeit einer Bindetasche berücksichtigt werden [27].

*Multi-objective*-Optimierungen versuchen, mehrere Optimierungsziele  $F=(f_1, \dots, f_m)$  gleichzeitig zu verfolgen. Diese Ziele sind an einen Vektor aus Parametern bzw. Entscheidungen  $X=(x_1, \dots, x_n)$  gekoppelt, welche wiederum abhängig sind von einer Menge  $k$  an Beschränkungen. Es wird dabei angenommen, dass alle Ziele über die Funktion  $f_i$  minimiert werden sollen. Die Erläuterung einer Dominanz-Selektion liefert nun folgender Satz: Für zwei Individuen, gegeben durch den Vektor  $X_1$  bzw.  $X_2$ , bei denen  $X_1$   $X_2$  dominiert, muss die Funktion  $f_i(X_1) \leq f_i(X_2)$  sein mit  $F(X_1) \neq F(X_2)$ . Gibt es mehrere gleich gute Individuen, so wird der Gewinner per Zufall ermittelt. Neben den Dominanz-selektierten Individuen werden noch die nicht-Dominanten beibehalten, welche nach jeder Bildung einer neuen Generation aktualisiert werden. Aus dieser Gruppe werden wiederum einige Individuen per Zufall ausgewählt und in die neu gebildete Generation eingefügt.

Nach einer vorher festgelegten Anzahl an erhobenen Generationen stoppt EMOS und somit der Optimierungsprozess. Abbildung 1.4 zeigt die schematische Darstellung des evolutionären Algorithmus mit seinen Einzelschritten.

Wie zu Beginn dieses Kapitels erwähnt, wurde zum Testen von EMOS die instabile F18K-Mutante von Villin Headpiece 36 (HP36) verwendet. Sowohl der HP36-Wildtyp als auch die destabilisierte F18K-Mutante wurden in einer MD (Molekulardynamik) simuliert und tatsächlich verlor die Mutante nach etwa 3 ns Simulation ihre strukturelle Stabilität und verblieb in einem hoch dynamischen, destabilisierten Zustand während der restlichen Simulationszeit von 7 ns. Das Ergebnis der Optimierung von F18K stellte die Rückführung der Mutante in den Ursprungszustand (HP36-Wildtyp) dar. Darin enthalten war überraschenderweise noch die zusätzliche G34L-Mutation für den Wildtyp. Es stellte sich heraus, dass diese Mutante eine erhöhte Stabilität gegenüber dem Wildtyp aufwies. Dies wurde dadurch festgestellt, dass während der gesamten Simulationszeit von 10 ns die RMSD der stabilen Mutante durchgehend bei etwa 0,25 nm lag, während sie für den Wildtyp konstant 0,35 nm und für die instabile F18K-Mutante bis zu 1 nm betrug. Die RMSD wurde dafür immer in Bezug zur Startstruktur berechnet.

Warum in der Natur dann nicht die stabilere Mutante auftaucht, lässt sich wie folgt erklären: Da HP36 nur einen Teil des Proteins Villin darstellt, könnte es sein, dass

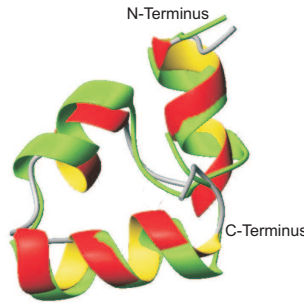


**Abbildung 1.4.:** Schematische Darstellung des evolutionären Algorithmus EMOS und dessen Einzelschritte. Verändert nach [27, 9].

die natürliche Evolution das Gesamtprotein optimiert hat und nicht nur HP36. Das G34 könnte einen biologischen Zweck erfüllen und darf deshalb nicht mutiert werden, ebenso kann es sein, dass die G34L-Mutante nicht mehr so gut wasserlöslich ist, was von Nachteil wäre, wenn es innerhalb einer Zelle funktionieren soll.

Drei Jahre nach Entwicklung von EMOS wurde parallel zu der gefundenen stabilen G34L-Mutante die dazu passende NMR-Struktur (pdb-ID: 2ppz) gelöst [36]. Diese zeigte eine sehr gute Übereinstimmung mit der über MD-Simulationen vorhergesagten Struktur (siehe Abb. 1.5), d. h. das Ziel der Wiederherstellung einer stabilen Tertiärstruktur wurde tatsächlich erfüllt und ließ sich mittels NMR-Spektroskopie nachweisen.

Der erste *in vitro*-Test der bioinformatischen Ergebnisse bestätigte somit die korrekte Funktion dieser Methode. Nicht nur die auffällig stabile G34L-Mutante befand



**Abbildung 1.5.:** Vergleich zwischen Modellstruktur (grün) und experimentell bestimmter Struktur der HP36-Mutante G34L (1ppz). Verändert nach [36].

sich in der Generation, die aus – laut EMOS – strukturell stabilen Individuen bestand, sondern auch weitere Mutanten. Die neun Mutanten, für welche stabile Konformationen vorhergesagt wurden, sind als multiples Alignment in Abbildung 1.6 zu sehen, ebenfalls eingefügt ist die Sequenz des HP36-Wildtyps 1vii.

Die Sequenzen können in zwei Cluster eingeteilt werden: in einen nicht-nativen (Sequenzen 1 bis 5) und in einen nativen (Sequenzen 6 bis 10). Als nicht nativ werden dabei alle Sequenzen gesehen, welche D4A-, W24E-, und/oder EKG32-34AYW-Mutationen tragen. Weiterhin ist auffällig, dass fast alle Mutanten die zu Beginn gesetzte, destabilisierende Mutation F18K beibehalten, vor allem die Mutanten aus dem nicht nativen Cluster. Die G34L-Mutante wird in diesem Aligment als r1g3i0 bezeichnet und gehört zu den nativen Mutanten.

nicht-nativ	1··r1g11i5	MY	SAEDFKAV	D	GMT	RS	AKANL	PL	E	KQQN	H	KKAYWLF					
	2··r1g12i6	MY	SAEDFKAV	D	GMT	RS	AKANL	PL	E	KQQN	D	KKAYWLF					
	3··r1g11i1	MY	SAEDFKAV	D	GMT	RS	AKANL	PL	E	KQQN	L	KKAYWLF					
	4··r1g8i2	MY	SAEDFKAV	F	GMT	RS	AKANL	PL	E	KQQN	L	KKAYWLF					
	5··r1g11i0	ML	SAEDFKAV	F	GMT	RS	AKANL	PL	E	KQQN	L	KKAYWLF					
nativ	6··1vii	ML	S	DEDFKAV	F	GMT	RS	A	NL	P	L	WKQQNL	KK	E	KGLF		
	7··r1g3i0	ML	S	DEDFKAV	F	GMT	RS	A	NL	P	L	WKQQNL	KK	E	KGLF		
	8··r2g4i0	ML	S	DEDFKAV	F	GMT	RS	A	NL	P	L	WKQQ	A	L	KK	E	KGLF
	9··r2g11i4	ML	S	DEDFKAV	F	GMT	RS	A	NL	P	L	WKQQ	C	L	KK	E	KGLF
	10r2g10i0	M	R	S	DEDFKAV	F	GMT	RS	AKANL	PL	E	KQQ	N	L	KK	MSG	L
				5	10	15	20	25	30	35							

**Abbildung 1.6.:** Sequenz-Alignment von stabilen Individuen, erstellt mit Prettyplot [37]. 1vii ist der HP36-Wildtyp, blau eingerahmt ist die in dieser Arbeit untersuchte 8x-Mutante r1g11i1. Verändert nach [36].



### 1.3. EMOS als neues Zusatzwerkzeug für Peptid-Design in der medizinischen Forschung

Im Laufe der Jahre wuchs das Interesse an künstlich erstellten, gleichzeitig auch biologisch aktiven Peptiden, da diese als mögliche Krankheitsbekämpfungsmittel Antibiotika und Antiviotika ergänzen können [38]. Ein besonders drängendes Problem ist die Zunahme der Resistenz von Krankheitserregern gegen die eingesetzten Anti-Infektiva. Einfache Infektionen können sich dadurch im schlimmsten Falle zu tödlichen Krankheiten entwickeln [39].

Einen natürlichen Ansatzpunkt – nicht nur zur Bekämpfung von Krankheiten, sondern auch zum Einsatz als Diagnosemittel – bietet daher die Forschung an Peptiden und Proteinen, die in jedem Organismus vorkommen und an allen lebenswichtigen Funktionen beteiligt sind, da sich überall in der Natur Peptide als Teil der natürlichen Abwehr, ob bei Pflanzen, Insekten, Fischen oder Säugetieren, finden [40]. Peptide nehmen somit in allen Organismen wichtige Funktionen ein, sie dienen als Hormone oder Neurotransmitter oder wirken im Fall von Defensinen wie Antibiotika [38].

Um gezielt Medikamente und Diagnosemittel auf Basis von Protein-Protein-Wechselwirkungen entwickeln zu können, ist es erstens notwendig, die Struktur beteiligter Proteine und Peptide zu kennen; zweitens wird mit einer möglichst präzisen Strukturvorhersage von medizinisch einsetzbaren Peptiden ein schnelles und flexibles Werkzeug für Peptiddesign geliefert. Somit müssen zwei Bereiche abgedeckt sein: Vorhersage bzw. Design und Aufklärung der Struktur von Peptiden.

Obwohl es bereits erfolgreiche Anwendungen für Peptide als Medikamente, wie z. B. dem Insulin oder dem HIV-Fusions-Inhibitor T20 oder Enfuvirtide [41, 42], gibt, existieren noch Faktoren, die die Wirkungsweise dieser Medikamente einschränken: Peptide denaturieren, werden abgebaut oder können aggregieren, sie sind große, meist hydrophile Moleküle mit flexiblen Bereichen, die biologische Membranen nur schwer überwinden können und bei denen aufgrund der Flexibilität die Bindung an ein Zielprotein entropisch ungünstig ist [43, 9]. Effektive Medikamente auf Peptid-Basis können daher nur hergestellt werden, wenn besagte Faktoren berücksichtigt werden.

Der evolutionäre Algorithmus EMOS stellt als ein Mehrziel-Optimierungsalgorithmus ein neues Werkzeug dar, welches zum Peptid-Design u. a. in der medizinischen Forschung eingesetzt werden kann. So können Peptide im Hinblick auf mehrere gewünschte Eigenschaften gleichzeitig optimiert werden, z. B. die Bindung an ein Zielprotein und Resistenz gegen Peptidasen [9]. Durch Einsatz dieses Algorithmus können aufwändige traditionelle *Screening*-Verfahren somit ergänzt werden, da gezielter nach Bindungspartnern gesucht werden kann.

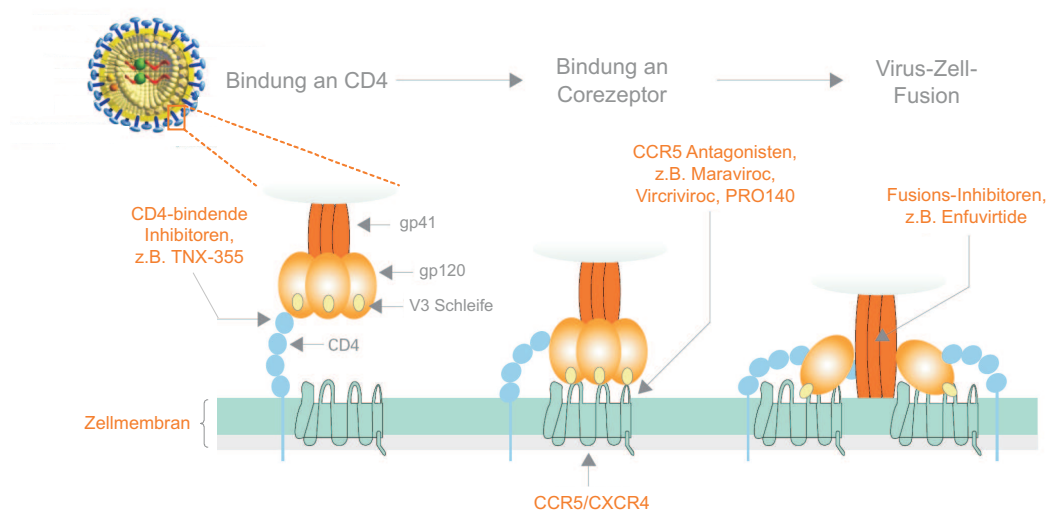
#### 1.3.1. Ein Beispiel für die Anwendung von EMOS

Proteine, die von menschlichen Zellen in die Blutbahn ausgeschieden werden oder an der Oberfläche als Rezeptoren für Hormone dienen, sind häufig chemisch verändert. Eine besondere chemische Modifizierung ist die Übertragung eines Sulfatrestes auf die Aminosäure Tyrosin. Derartige Sulfotyrosine wirken als „molekulare Klebstoffe“ bei



der Wechselwirkung des Proteins mit anderen Molekülen. Sulfotyrosine in Hormonen und Chemokinen sind für die Bindung an ihre Rezeptoren notwendig. Sulfatierte Tyrosine finden auch ihren Einsatz in der Blutgerinnungskaskade oder ermöglichen die Anheftung von Leukozyten an die Epithelschicht in der Blutbahn, wo diese bakterielle Infektionen bekämpfen können. Viren wie HIV nutzen Sulfotyrosine an der Oberfläche von menschlichen Zellrezeptoren, um in die Zelle einzudringen und sie zu infizieren [44, 45].

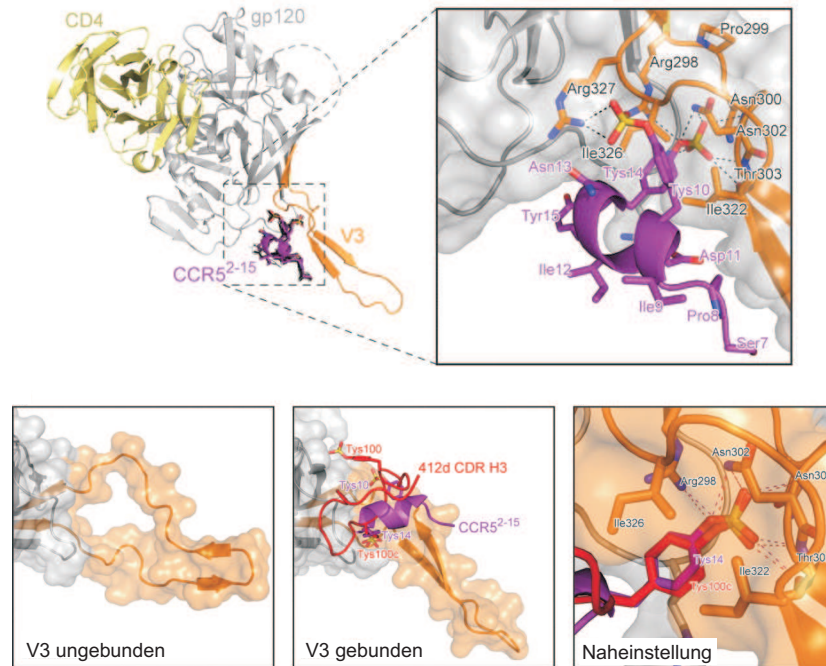
Eine anschauliche Erklärung dafür, wie diese Idee der Generierung künstlicher Peptide praktisch angewendet werden kann, liefert folgendes Beispiel: Ein HI-Virus muss, um sein genetisches Material in eine menschliche Zelle übertragen – und sie damit infizieren – zu können, an diese binden. Dies geschieht durch Andocken des viralen Hüllproteins gp120 an die menschlichen Oberflächenrezeptoren CD4 und CCR5 oder CXCR4 (siehe Abb. 1.7). Jedoch erst nach Anbinden an CCR5 bzw. CXCR4 kann das Virus seine RNA in die menschliche Zelle einschleusen. Kann diese Anbindung verhindert werden, so würde eine Infektion der Zelle ausbleiben.



**Abbildung 1.7.:** Schematische Darstellung des HIV-Infektionsmechanismus. gp120, ein Glycoprotein an der Zelloberfläche von HIV, bindet an den menschlichen Rezeptor CD4 und anschließend an CCR5 oder CXCR4. Bereits erhältliche Medikamente setzen an verschiedenen Punkten an, um eine Infektion der Zelle mit HIV zu vermeiden. Abbildung erstellt nach [46, 47, 48].

Im März 2003 wurde das anti-HIV Peptid Enfuvirtide (Fuzeon) von der Firma Roche entwickelt [43]. Dieses Medikament ist die natürlich vorkommende Interface-Sequenz (36 Aminosäuren) des Glycoproteins gp41. Bindet gp120 an den humanen Rezeptor CD4, so ändert sich die Konformation von gp120 und das virale gp41 wird exponiert, welches den Virus dabei unterstützt, mit der Membran der menschlichen Zelle zu fusionieren, um dann sein genetisches Material einschleusen zu können. Enfuvirtide inhibiert den letzten viralen Fusionsschritt, indem es an gp41 bindet (siehe Abb. 1.7).

Enfuvirtide stellt nur eines der HIV-Therapeutika dar, welche mittlerweile auf den Markt gekommen sind. So gibt es u. a. auch CD4-bindende Inhibitoren wie TNX-355 [49] und CCR5-Antagonisten wie Maraviroc [50] (siehe Abb. 1.7). Die Entwicklung solcher Medikamente ist sehr aufwändig und dauert Jahrzehnte. Gleichzeitig muss bedacht werden, dass starke Nebenwirkungen auftreten. So kann z. B. die Blockierung des humanen CD4-Rezeptors das humane Immunsystem beeinflussen, da dort auch in erstaunlich ähnlicher Weise wie gp120 das Klasse II-MHC-Protein bindet [51, 52].



**Abbildung 1.8.:** Strukturelle Darstellung der Bindung des N-Terminus von CCR5 an gp120 (oben), die strukturelle Umlagerung der V3-Schleife (unten links, Mitte) nach Anlagerung sowohl von CCR5 als auch vom Antikörper 412d CDR H3. Verändert nach [53].

Da der gp120/CD4-Komplex noch an einen der beiden Corezeptoren CCR5/CXCR4 bindet, bietet letzterer einen weiteren Angriffspunkt in der Entwicklung von HIV-Therapeutika. TAK-779 wurde Ende der 1990er Jahre entwickelt und ist ein nicht-peptidäres Medikament, welches in eine Bindetasche bindet, die durch die transmembranen Domänen des CCR5 ausgebildet wird. Es wird eine Konformationsänderung in CCR5 induziert, so dass eine Bindung an das gp120 nicht mehr möglich ist. Die natürliche Funktion von CCR5 wird dabei nicht beeinträchtigt [54]. Allerdings zeigt sich, dass nach einer bestimmten Zeit Resistenzen gegenüber TAK-779 auftreten [55].

Der gp120/CD4-Komplex bindet an sulfatierte Tyrosine am N-Terminus des CCR5 [56, 57]. Abbildung 1.8 zeigt, wie die sulfatierten Reste Tys10 und Tys14 des N-Terminus von CCR5 zur Bindung an die V3-Schleife des viralen gp120 beitragen (oben) und die V3-Schleife nach Anbinden von CCR5 seine Konformation ändert (unten). Ebenfalls dargestellt ist in rot der Antikörper 412d CDR H3, der ähnlich wie CCR5

über das Tys100c an gp120 bindet (unten Mitte, unten rechts). So können auch an dieser Stelle Therapeutika angreifen. Allerdings sind diese wiederum teilweise schwer löslich in wässriger Umgebung [58] und sie sind der schnellen Mutationsrate mit etwa  $3 \cdot 10^{-5}$  pro Nukleotidbase pro Replikationszyklus ( $10^{10}$ - $10^{12}$  Viren pro Tag) [59, 60] des HI-Virus unterworfen, da sie am viralen Protein angreifen.

So gibt es zahlreiche Herausforderungen wie auftretende Resistenzen, starke Nebenwirkungen und schlechte Löslichkeit von Peptiden als Medikamente, die besagen, dass eine Verbesserung der Therapeutika benötigt wird. Genau an diesem Punkt setzt EMOS an und versucht, *in silico* bereits bekannte oder auch neue Peptide, die als Medikamente eingesetzt werden können, zu optimieren und so die Verbesserung zu beschleunigen und den Optimierungsaufwand so minimal wie möglich zu halten. Langfristig könnten so Peptide designed werden, welche auf den menschlichen Körper hin angepasst werden würden. So könnten relevante Parameter wie beispielsweise Resistenz gegenüber Peptidasen trotz möglichst geringer Größe des Peptids und eine hohe Affinität zum potentiellen Bindepartner sowie eine gute Löslichkeit in wässriger Umgebung über diesen Mehrziel-Optimierungsalgorithmus gleichzeitig berücksichtigt werden, um schnell und effizient an das gewünschte medizinische Therapeutikum zu gelangen. Je genauer bekannt ist, wie in dem hier beschriebenen Fall HIV an die menschliche Zelle bindet, also welche Aminosäuren konkret an der Bindung beteiligt sind und wie das Bindeinterface aussieht, desto präziser kann EMOS arbeiten.

## 1.4. Zielsetzung

Natürliche Peptide wie Oxytocin, Cyclosporine und Insulin sowie synthetisch hergestellte Peptide wie Fuzeon und Integrilin gehören zu den zugelassenen Peptid-basierten Medikamenten. Peptid-basierte Medikamente besitzen verglichen mit Therapeutika kleiner molekularer Größe einige Vorteile, sie sind spezifischer, nicht so toxisch und werden weniger im Gewebe akkumuliert. Natürlich bergen sie auch einige Nachteile, da sie relativ schnell vom Körper abgebaut werden, weniger stabil und teurer in der Synthese sind [41].

Peptide besitzen großes Potential im Einsatz als effektive Medikamente, so gibt es bislang mehr als 40 Peptid-basierte Therapeutika auf dem Markt. 270 Peptide befinden sich bereits in der klinischen Phase, etwa 400 Peptide in der vorklinischen Testphase [61, 62, 63]. Zeitaufwendige, kostenträchtige und viel Material verbrauchende Synthesen stellen jedoch immer noch ein Problem für Hochdurchsatz-Verfahren dar, in denen Peptide parallel auf ihre Wirkung und Nützlichkeit hin untersucht werden. Daher werden neue Verfahren benötigt, die sowohl Zeitaufwand als auch Kosten und Verbrauch von Chemikalien und anderen Materialien reduzieren [41]. Da die Peptidsynthese teuer ist und stets eine große Anzahl an verschiedenen Peptiden über Hochdurchsatz-Verfahren getestet werden muss, um das als potentiell Therapeutikum einsetzbare Peptid zu finden, sollte dieses ineffiziente Verfahren durch neue Methoden ersetzt bzw. ergänzt werden [64, 65].

Struktur-basiertes *Drug-Design* am Computer kann an dem Punkt als effektive, neue Methode eingesetzt werden. Fortschritte in der Strukturaufklärung von Proteinen und Peptiden über Röntgenkristallographie und NMR-Spektroskopie ergänzen die Strategie des bioinformatischen *Drug-Designs* dadurch, dass bereits bekannte Strukturen als Vorlage für die Bioinformatik dienen können. So determinierte Bindetaschen führen z. B. zu der Entwicklung von einer Vielzahl an *Docking*-Strategien.

Es existieren zahlreiche Berichte über Entdeckung an Therapeutika von Bibliotheken aus kleinen Liganden. Allerdings ist es immer noch schwierig, die *Docking*-Energien aller Peptide zu berechnen, da diese sehr unterschiedlich sind [66, 67].

Effektive Medikamente auf Peptid-Basis müssen daher weiterhin gefunden und optimiert werden. Es existieren bereits jetzt Berichte, die beschreiben, dass erfolgreich Peptide hinsichtlich eines Parameters optimiert wurden [68, 69]. Dennoch sollte eine Optimierung darauf abzielen, mehrere Parameter parallel verbessern zu können, wie z. B. die Affinität eines Liganden zu seiner Bindetasche zu erhöhen und gleichzeitig die Löslichkeit des Peptids in wässriger Umgebung beizubehalten oder zu optimieren.

Ein Ansatzpunkt, um optimale Peptidsequenzen zu finden, ist die Verwendung von evolutionären Algorithmen, welche hinsichtlich mehrerer Parameter gleichzeitig optimieren können. Dies ist bislang noch kein gängig genutztes Konzept, um *in silico* therapeutisch wirksame Peptide zu erstellen. So existieren nur wenige Belege aus der Bioinformatik, in denen diese multiobjektiven Lösungswege genutzt werden, gerade im Bereich des kombinatorischen Designs von Peptid-Bibliotheken [70, 71, 72].

Der im Jahre 2005 entwickelte Algorithmus EMOS (*Evolutionary Multiobjective Optimization System*) bildet als Neuentwicklung eines solchen Verfahrens, der multiobjektiv ist, den Ansatzpunkt für die Identifizierung und Optimierung von Peptiden, welche

als potentielle Therapeutika eingesetzt werden könnten. Da ein solcher Algorithmus sehr komplexe Aufgabenstellungen lösen muss und gleichzeitig dabei wenig bzw. keine Fehler machen sollte, ist es notwendig, ihn ausreichend zu testen und kontinuierlich zu optimieren.

Um EMOS testen und optimieren zu können, wird ein Peptid als Modellsystem benötigt, welches bereits ausreichend charakterisiert sein sollte. Dafür bietet sich Villin Headpiece 36 an, da es klein ist und sowohl über bioinformatische Methoden als auch über NMR-Spektroskopie bereits sehr gut untersucht ist (siehe Kap. 1.1). Obwohl EMOS über die NMR-spektroskopische Strukturaufklärung der nativen G34L-Mutante auch *in vitro* als wahrscheinlich korrekt funktionierend bestätigt wurde, fiel doch eine besondere Aufmerksamkeit auf den Cluster von nicht nativen Mutanten, welche sich mit bis zu 9 Mutationen gegenüber dem Wildtyp unterschieden und somit deutlich mehr Aminosäuren ausgetauscht hatten als die Einzelmutante G34L (siehe Kap. 1.2).

Die Effizienz von EMOS sollte anhand einer nicht nativen Mutante hinterfragt und charakterisiert werden. Eine dieser nicht nativen Mutanten ist r1g11i1 (siehe Abb. 1.6, blau umrahmt) und wurde für dieses Vorhaben ausgewählt.

Die Struktur der Mutante r1g11i1 wurde – wie auch vorangehend die für die G34L-Mutante – über eine MD-Simulation vorhergesagt und lautet new2. Die stabile Struktur als Endprodukt der Durchläufe des evolutionären Algorithmus dient somit u. a. als Bewertungskriterium, ob EMOS korrekt funktioniert. Sollte dies der Fall sein, so kann die Bestätigung dieser Struktur *in vitro* die Funktionalität dieses Algorithmus untermauern. Sollte das Gegenteil eintreten, so kann das Ergebnis dazu dienen, EMOS zu optimieren.

Das Ziel dieser Arbeit besteht darin, die achtfache Villin Headpiece-Mutante (r1g11i1) in ihrer Struktur und Dynamik zu charakterisieren und mit der Vorhersage new2 zu vergleichen. Das Peptid wird im weiteren Verlauf dieser Arbeit als mut8 (Achtfachmutante des HP36) bezeichnet, um es von der vorhergesagten Primärsequenz r1g11i1 und der bioinformatischen Vorhersage new2 zu unterscheiden. Zur Charakterisierung von mut8 soll das chemisch synthetisierte Peptid über NMR-Spektroskopie untersucht werden. In einem weiteren Schritt soll mut8 kloniert,  $^{15}\text{N}$ -markiert exprimiert und gereinigt werden, um wiederum über NMR-spektroskopische Verfahren zusätzliche Informationen zur Charakterisierung des Peptids zu gewinnen. Die so erworbenen Kenntnisse sollen dann an die Arbeitsgruppe des Entwicklers, Prof. Daniel Hoffmann (Universität Duisburg-Essen, ZMB, Bioinformatik), weitergegeben werden, um die Effizienz von EMOS bestätigen oder verbessern zu können.

Langfristig soll EMOS so optimiert werden, dass es als effizientes und korrekt funktionierendes Werkzeug zum Design von Peptiden eingesetzt werden kann.

## 2. Materialien und Methoden

### 2.1. Materialien

#### 2.1.1. Chemikalien

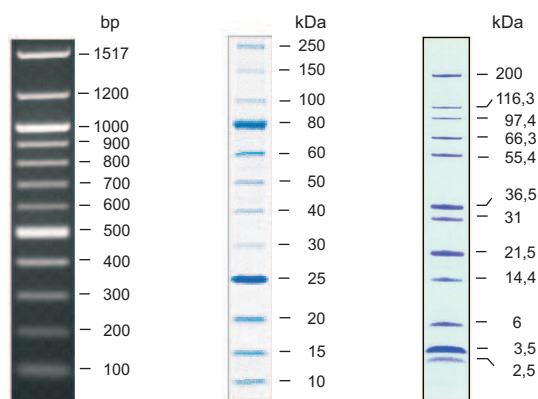
Allgemein gebräuchliche Laborchemikalien wurden, falls nicht anders vermerkt, von den Firmen Roth (Karlsruhe), Merck (Darmstadt) oder Sigma-Aldrich (Seelze) bezogen.

AA/Bis Rotiphorese Gel 30	Roth, Karlsruhe
Agarose	Roth, Karlsruhe
$^{15}\text{N}$ -Ammoniumchlorid ( $^{15}\text{NH}_4\text{Cl}$ )	Spectra, Cat. No. 5300, Lot. No. 1385
Ammoniumperoxodisulfat (APS)	Roth, Karlsruhe
Borsäure ( $\text{H}_3\text{BO}_3$ )	Roth, Karlsruhe
Calciumchlorid ( $\text{CaCl}_2$ )	Roth, Karlsruhe
Cobaltchlorid-Hexahydrat ( $\text{CoCl}_2 \cdot 7\text{H}_2\text{O}$ )	Roth, Karlsruhe
Coomassie G250	Fluka, Seelze
Deuteriumoxid ( $\text{D}_2\text{O}$ ), 99,9 %	Deutero GmbH, Kastellaun
Dinatriumhydrogenphosphat Heptahydrat ( $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ )	Roth, Karlsruhe
1,4-Dithiothreitol (DTT)	Biomol, Hamburg
2,2-Dimethyl-2-silapentan-5-sulfonsäure (DSS)	Sigma-Aldrich, Seelze
Eisen-III-Citrat (Fe-III-Citrat)	Sigma-Aldrich, Seelze
Ethidiumbromid	Roth, Karlsruhe
Hefeextrakt	Roth, Karlsruhe
4-2-Hydroxyethylpiperazinoethansulfonsäure (Hepes)	AppliChem, Darmstadt
Isopropylthiogalactopyranosid (IPTG)	Gerbu, Gaiberg
Kaliumdihydrogenphosphat ( $\text{KH}_2\text{PO}_4$ )	Roth, Karlsruhe
Kanamycinsulfat	Gerbu/Roth, Karlsruhe
Kupferchlorid-Dihydrat ( $\text{CuCl}_2$ )	Roth, Karlsruhe
Magnesiumsulfat ( $\text{MgSO}_4$ )	Fluka, Seelze
Manganchlorid ( $\text{MnCl}_2$ )	Roth, Karlsruhe
Natriumazid ( $\text{NaN}_3$ )	Roth, Karlsruhe
Natriumdodecylsulfat (SDS)	Roth, Karlsruhe
Natriummolybdat-Dihydrat ( $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ )	Roth, Karlsruhe
Natriumselenit-Pentahydrat ( $\text{Na}_2\text{SeO}_3$ )	Fluka, Seelze

Nickel-II-Chlorid Hexahydrat ( $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$ )	Roth, Karlsruhe
Orange G	Roth, Karlsruhe
Pepton	Roth, Karlsruhe
Phenylmethanolsulfonylfluorid (PMSF)	Roth, Karlsruhe
Thiaminchlorid (Vitamin B <sub>1</sub> )	Merck, Darmstadt
Tris-hydroxymethylaminomethan (Tris)	Roth, Karlsruhe
Zinksulfat-Heptahydrat ( $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ )	Roth, Karlsruhe

### 2.1.2. DNA- und Protein-Marker

Als DNA-Marker wurde der 100 bp-Marker (New England Biolabs) verwendet (siehe Abb. 2.1, links). Als Protein-Marker wurden der 10-250 kDa Protein-Marker (New England Biolabs) und der Mark12 *Unstained Standard* (Invitrogen) verwendet (siehe Abb. 2.1, Mitte und rechts).



**Abbildung 2.1.:** 100 bp-Marker (New England Biolabs) (links), Protein-Marker 10-250 kDa (New England Biolabs) (Mitte) und Mark12 *Unstained Standard* (Invitrogen) (rechts).

### 2.1.3. Puffer und Lösungen

Die Puffer des gebrauchsfertigen Kits für die Plasmid-Minipräparation werden, wie vom Hersteller geliefert, benutzt, ihre Zusammensetzung ist dabei dem Beipackzettel zu entnehmen.

Bradford-Reagens:	100 mg Coomassie Brilliant Blau G250
	50 ml 95 % Ethanol
	100 ml konz. Phosphorsäure
	ad 200 ml Aqua bidest

## 2. Materialien und Methoden

---

Coomassie-Färbe-Lösung:	1,25 g Coomassie Brilliant Blau G250 227 ml Ethanol 46 ml Essigsäure ad 227 ml Aqua bidest
DNA-Probenpuffer (6 x):	40 % (w/v) Saccharose 0,5 % Orange G
Elutionspuffer:	50 mM Tris-HCl; pH 7,1 50 mM NaCl 10 mM Glutathion 1 mM PMSF 250 $\mu$ M EDTA  → nach Zugabe von Glutathion pH erneut einstellen!
Entfärbe-Lösung:	75 ml Ethanol 50 ml Essigsäure 875 ml Aqua bidest
Gelfiltrationspuffer:	50 mM KPi; pH 6,5 150 mM KCl 1 mM PMSF 250 $\mu$ M EDTA
Hochsalzpuffer:	50 mM Tris-HCl; pH 7,1 400 mM NaCl 1 mM PMSF 250 $\mu$ M EDTA
KPi-Puffer:	50 mM KPi; pH 6,5
Protein-Probenpuffer (2 x):	125 mM Tris-HCl; pH 6,8 50 % (w/v) Glycerin 10 % (w/v) SDS 14 mM $\beta$ -Mercaptoethanol 0,01 % Bromphenolblau
Sammelgelpuffer:	0,5 M Tris; pH 6,8
SDS-PAGE-Laufpuffer (10 x):	250 mM Tris; pH 8,3 1 % (w/v) SDS 2,5 M Glycin



TAE-Puffer (50 x):	2 M Tris-Acetat; pH 8,3 50 mM EDTA
Trenngelpuffer:	1,5 M Tris; pH 8,0
TS2 (50x):	700 mg $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 30 mg $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ 300 mg $\text{H}_3\text{BO}_3$ 200 mg $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ 200 mg $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$ 10 mg $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ 900 mg $\text{Na}_2\text{MoO}_4$ 20 mg $\text{Na}_2\text{SeO}_3$ ad 1000 ml Aqua bidest  → nicht autoklavieren, sondern nur steril filtrieren!

Alle biochemischen und molekularbiologischen Arbeiten wurden mit Aqua bidest aus einem Milli-Q Biocel Reinstwasseraufbereitungssystem (Millipore, Billerica, USA) durchgeführt.

### 2.1.4. Nährmedien und Antibiotika

LB-Medium: 10 g Pepton  
5 g Hefeextrakt  
10 g NaCl; pH 7  
50  $\mu\text{g}/\mu\text{l}$  Kanamycin  
ad 1 l Aqua bidest

LB-Agarose: 15 g Agarose  
ad 1 l LB-Medium

Die Herstellung von Agar-Platten erfolgt durch Zugabe einer bestimmten Menge (w/v) an Agar-Agar in das LB-Medium, welches anschließend autoklaviert wird. Nach dem Abkühlen auf etwa 50 °C werden weitere sterile Komponenten wie Antibiotika hinzugefügt, anschließend werden je 20 ml des noch flüssigen Mediums in Petrischalen gegossen. Gelagert werden die Platten bei 4 °C.

M9-Medium: 51,46 g Na<sub>2</sub>HPO<sub>4</sub>·7H<sub>2</sub>O  
12 g KH<sub>2</sub>PO<sub>4</sub>  
4 g <sup>15</sup>NH<sub>4</sub>Cl  
  
8 ml MgSO<sub>4</sub> (1 M)  
8 ml TS2 (50x)  
9,8 mg Fe-III-Citrat  
400 µl Vitamin B<sub>1</sub> (0,5 %)  
200 mg Kanamycin  
400 µl CaCl<sub>2</sub> (1 M)

Als Antibiotikum wurde Kanamycin eingesetzt und diente zur Selektion positiver Transformanten nach der Klonierung (siehe 2.3.1) und während der Expression (siehe 2.3.3).

Kanamycin liegt in Pulverform vor, die sachgemäße Lagerung erfolgt bei 4 °C. Für den Gebrauch wird das Antibiotikum in einer Konzentration von 50 mg/ml in steril filtriertem Wasser gelöst und direkt verwendet oder bei -20 °C gelagert.

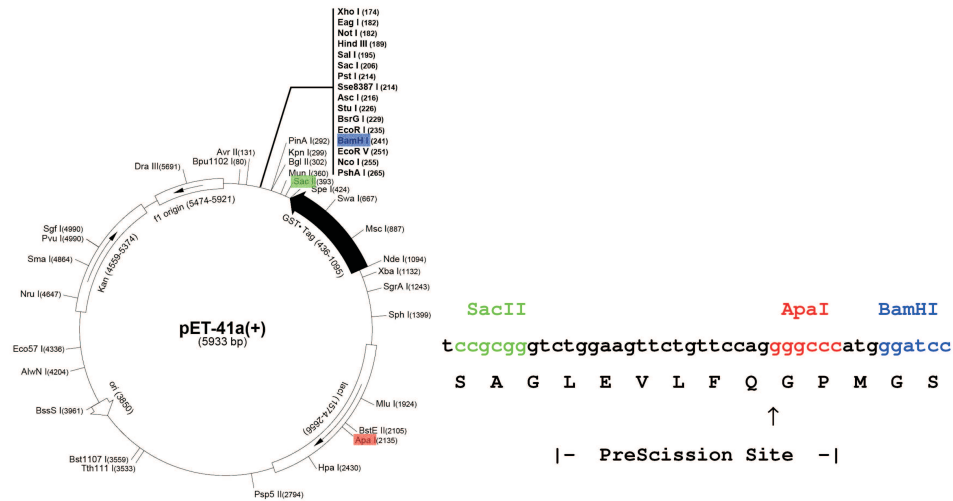
### 2.1.5. Bakterienstämme

Verwendet wurde zur Anzucht für eine Plasmid-Minipräparation und zur Überprüfung der korrekten Klonierung der Stamm *Escherichia coli* XL10 Gold. Zur Expression des Peptids wurde *Escherichia coli* Rosetta (DE3) verwendet.

Stamm	Genotyp
<i>E. coli</i> XL10 Gold	Tet <sup>r</sup> Δ( <i>mcrA</i> )183 Δ( <i>mcrCB-hsdSMR-mrr</i> )173 <i>endA1 supE44 thi-1 recA1 gyrA96 relA1 lac</i> Hte [F' <i>proAB lacI<sup>q</sup>ZΔM15 Tn10</i> (Tet <sup>r</sup> ) Amy Cam <sup>r</sup> ]
<i>E. coli</i> Rosetta (DE3)	F <sup>-</sup> <i>ompT hsdSB(r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) gal dcm lacY1</i> (DE3) pRARE6 (CmR)

### 2.1.6. Plasmid

Die nachfolgende Klonierung wurde ausgehend vom Vektor pET41b\_mod erstellt. Dieser lag als ein bereits im Voraus institutsintern modifizierter pET41-Vektor vor (Merckbiosciences), nachdem nach Deletion der ApaI-Schnittstelle eine PreScission-Protease-Erkennungssequenz inklusive ApaI-Schnittstelle innerhalb der *multiple cloning site* (MCS) über SacII und NcoI einkloniert wurde (siehe 2.2).



**Abbildung 2.2.:** Schematische Darstellung des pET41b\_mod-Vektors (links) mit einklonierter PreScission-Protease-Erkennungssequenz (rechts) [73].

### 2.1.7. Enzyme

Klonierung:	Pfu-Polymerase, 5 U/ $\mu$ l (Metabion, Martinsried)
	<i>Fast Digestion</i> -Restriktionsendonukleasen
	XhoI und ApaI (Fermentas, St. Leon-Rot)
	Ligase, 150 U/ $\mu$ l (Metabion, Martinsried)
Abschneiden des GST-tags:	PreScission-Protease (GE Healthcare, Freiburg)
Zellwandlyse:	Lysozym (Fluka, München)

### 2.1.8. Verbrauchsmaterialien

NMR-Röhrchen	NMR Sample Tubes AP5-800-7 (Ampol)
Sterilfilter	Express PLUS 0,22 $\mu$ m (Millipore)
UV-Küvetten	Halbmikro-Einmalküvetten, ratiolab

### 2.1.9. Geräte

Folgende Auflistung gibt einen Überblick über alle verwendeten Geräte. Kleingeräte, die hier nicht aufgeführt wurden, entsprechen der Laborstandardausrüstung.

Gerät	Modell	Hersteller
Autoklav	H+P Varioklav 25T	Oberschleißheim
Brutschrank	Inkubationsschrank	Binder, USA
FPLC	Äkta FPLC P920	GE Healthcare, Solingen
Inkubatoren	Thermoschüttler HT	Infors, Bottmingen
	Thermomixer <i>comfort</i>	Eppendorf, Hamburg
Konzentratoren	Amicon Ultra, 10000 MWCO	Millipore, Schwalbach
	Vivaspin 2000 MWCO, 15R	Sartorius, Göttingen
Magnetrührgeräte	Assistant RM5	Heidolph Instruments, Schwabach
Lyophylle	Gefriertrocknungsanlage alpga 2-4 LD plus	Christ, Osterode
pH-Meter	766 Calimatic mit SE100 Elektrode	Knick, Berlin
Photometer	Bio Photometer	Eppendorf, Hamburg
Pumpen	Hochvakuumpumpe RZ 2.5	Vacuubrand GmbH, Wertheim
PCR-Gerät	Thermocycler T3000	Biometra, Göttingen
Rotoren	JLA 9.1000 Ti45	Beckman, USA
		Beckmann, USA
Spektrometer	Ultrashield Plus 400 NMR-Spektrometer	Bruker, BioSpin, Fällanden
Sterilbank	HERA safe	Thermo Scientific, Langenselbold
UV-Tisch	BioDoc Analyze Ti5	Biometra, Göttingen
	UV-Flächenstrahler N36-M	Benda, Wiesloch
Wasseraufbereitung	Milli-Q Biocel Reinstwasseraufbereitungssystem	Millipore, Bedford, USA
Zellaufschlussanlage	Microfluidics Microfluidizer M-110L	Newton, Massachusetts
Zentrifugen	Tisch-Kühlzentrifuge 5415R	Eppendorf, Hamburg
	Tischzentrifuge MiniSpin Plus	Eppendorf, Hamburg
	Centrifuge 5810R	Eppendorf, Hamburg
	Avanti JE	Beckman, Palo Alto
	Optima LE-80K Ultrazentrifuge	Beckman, Palo Alto

## 2.1.10. Software

Datenanalyse:

Qti-Plot

[<http://soft.proindependent.com/qtiplot.html>]

Origin 8.1

[<http://www.originlab.com/>]

Ramachandranplots:

RamPage (Cambridge) [1]

Sekundärstrukturanteile:

MoleMan (Uppsala) [74]

Sequenzalignment:

ClustalW (Gibson, Thompson, Higgins) [2]

JalView (University of Dundee) [3, 75]

Spektrenaufnahme/Prozessierung:

TopSpin 2.0 (Bruker, Karlsruhe)

Strukturberechnung:

Xplor-NIH 2.18 [76, 77]

Visualisierung:

PyMol 0.99 rc6 [[www.pymol.org](http://www.pymol.org)]

Zuordnung und Integration:

AUREMOL (Bruker, Rheinstetten) [78]

SPARKY (Goddard *et al.*, San Francisco) [79]

## 2.2. Peptidsynthese

Das in dieser Arbeit verwendete Peptid mut8 wurde von der Firma Bachem U.K. LTD synthetisiert. Die Aminosäuresequenz lautet:

N-term–MYSAEDFKAVDGMTRSAKANLPLEKQQNLKKAYWLF–C-term

Das 36 Aminosäuren lange Peptid wurde über Festphasensynthese (SPPS) hergestellt. Dafür werden vom C-Terminus ausgehend die geschützten Aminosäuren nach und nach an eine feste Matrix, einem nicht löslichen Polymer auf Polystyrol-Basis, gebunden. Jede Aminosäurenkopplung besteht aus drei Schritten: Abspaltung der  $\alpha$ -Amino-Schutzgruppe, Koppeln einer Fmoc-geschützten Aminosäure, Waschen mit organischen Lösungsmitteln. Am Ende der Synthese wird das fertige Peptide unter sauren Bedingungen von der Matrix abgespalten. Zur Reinigung des hergestellten Peptids wird eine *reversed phase* (RP)-HPLC durchgeführt. Die dazu verwendeten Puffer sind Wasser mit 0,1 % Trifluoressigsäure (TFA) als Puffer A und Acetonitril mit 0,1 % TFA als Puffer B. Während der Elution werden die Peptid-haltigen Fraktionen gesammelt, auf ihre Reinheit hin analysiert und schließlich lyophilisiert. Die Reinheit wird definiert als prozentualer Anteil des Peptids verglichen mit der Gesamtmenge an Verbindungen, die bei 210 nm (TFA) oder bei 220 nm (Triethylammoniumphosphat, TEAP) absorbieren. Die Reinheit wird für dieses Peptid als  $\geq 95$  % angegeben [80]. Detailliertere Angaben zur Peptidsynthese sind aufgrund von Betriebsinterna nicht möglich.

## 2.3. Mikro- und Molekularbiologische Methoden

### 2.3.1. Klonierung von mut8

Das synthetische Gen (*E. coli* Codon usage, codierend für mut8) folgender Sequenz wurde als Lyophilisat von der Firma Microsynth bezogen.

Oberer Strang:

5'- CAT GTA TAG CGC AGA AGA TTT TAA AGC AGT TGA TGG TAT GAC  
CCG TAG CGC AAA AGC AAA TCT GCC GCT GGA AAA ACA GCA GAA TCT  
GAA AAA AGC ATA TTG GCT GTT TTA ATC-3'

Unterer Strang (revers komplementär):

5'-TCG AGA TTA AAA CAG CCA ATA TGC TTT TTT CAG ATT CTG CTG  
TTT TTC CAG CGG CAG ATT TGC TTT TGC GCT ACG GGT CAT ACC ATC  
AAC TGC TTT AAA ATC TTC TGC GCT ATA CAT GGG CC-3'

Vor der eigentlichen Klonierung wurde eine elektronische Kontroll-Klonierung durchgeführt, d. h. per Computer wurde das für mut8 codierende Gen über die Schnittstellen in den Zielvektor eingefügt (analog zu Kap. 2.3.1). Diese elektronische Klonierung

diente später zur Kontrolle der *in vitro* durchgeführten Klonierung.

### Annealing des synthetischen Gens und PCR-Reaktion

Jeder der beiden Lyophilisate wurden in Aqua bidest gelöst, je ein Aliquot entnommen und in einem 1:1-Verhältnis (w/w) in einer Endkonzentration von 300 pmol/ $\mu$ l gemischt. Für ein Anlagern beider Sequenzen wurde das Gemisch in einem auf 95 °C erhitzten Heizblock auf selbige Temperatur erhitzt, der Heizblock ausgeschaltet und der Ansatz – im Heizblock belassend – auf Raumtemperatur abgekühlt.

Die Probe wurde auf 0,3 pmol/ $\mu$ l verdünnt und 2,3  $\mu$ l (Endkonz.: 0,007 pmol/ $\mu$ l) wurden in ein PCR-Reaktionsgefäß gegeben. Hinzugefügt wurden 11,5  $\mu$ l Pfu-Puffer (10x), 0,23 mM dNTPs, 10 pmol/ $\mu$ l je Primer 646\_F bzw. 647\_R (siehe unten) und 0,09 u/ $\mu$ l Pfu-Polymerase. Der Ansatz wurde auf 100  $\mu$ l mit Aqua bidest aufgefüllt.

Die zur Klonierung von mut8 verwendeten Primer 646\_F und 647\_R wurden von der Firma Metabion in einer Konzentration von 100  $\mu$ M bezogen:

Forward-Primer (646\_F):

5'-CAC ACA CAG GGC CC ATG TAT AGC GCA GAA G-3'

Reverse-Primer (647\_R):

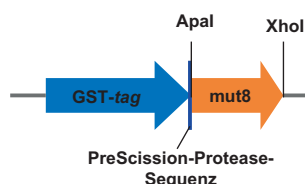
3'-GGA AGG CTC GAG ATT AAA ACA GCC AAT ATG CTT TTT TC-5'

Die PCR wurde folgendermaßen durchgeführt: Initiale Denaturierung bei 94 °C für 2 min, Denaturierung bei 94 °C für 30 s, gefolgt von einer Anlagerungsphase bei 70 °C für 30 s und einer Elongationsphase bei 72 °C für 1 min (35x). Die finale Elongationsphase bei 72 °C für 10 min bildete den Abschluss. Das PCR-Produkt wurde auf ein 1,5 %iges Agarosegel aufgetragen, die Bande bei etwa 140 bp ausgeschnitten und über das Gelextraktions/PCR-Kit (NucleoSpin ExtractII, Macherey-Nagel) gereinigt.

### Restriktion und Ligation

Das PCR-Produkt wurde ebenso wie der Vektor pET41b\_mod mit ApaI und XhoI restringiert. Dafür wurden 0,5  $\mu$ g des PCR-Produktes bzw. 1  $\mu$ g des Zielvektors mit 2  $\mu$ l *Fast Digestion Buffer* versehen und je 1  $\mu$ l ApaI bzw. XhoI hinzugegeben. Der Restriktionsansatz wurde auf 20  $\mu$ l mit Aqua bidest aufgefüllt und alles für 10 min bei 37 °C inkubiert. Beide Ansätze wurden über das Gelextraktions/PCR-Kit (NucleoSpin ExtractII, Macherey-Nagel) gereinigt. Nach der Restriktion wurden Vektor und PCR-Produkt in einem Verhältnis von 5:1 (Amplifikat:Vektor) ligiert. Dafür wurden die 0,5  $\mu$ g des PCR-Produktes mit 0,1  $\mu$ g Vektor gemischt und 1 mM ATP, 2  $\mu$ l Ligase-Puffer sowie 0,5  $\mu$ l Ligase hinzugegeben. Der Ansatz wurde auf 20  $\mu$ l aufgefüllt und die Ligation bei 17 °C über Nacht durchgeführt.

Nach einer erfolgreichen Klonierung ergibt sich das in Abb. 2.3 schematisch dargestellte Konstrukt GST\_mut8 (dargestellt als Ausschnitt aus dem fertig klonierten



**Abbildung 2.3.:** DNA-Konstrukt zur Expression von *mut8*. Der GST-tag (blau) befindet sich N-terminal von *mut8* (orange) und kann nach der Expression an der PreScission-Protease-Sequenz abgeschnitten werden.

Vektor pET41b\_mod\_GST\_ mut8) mit den Schnittstellen *ApaI* und *XhoI* sowie der PreScission-Protease-Sequenz zum Entfernen des GST-tags nach einer Reinigung.

### 2.3.2. Überprüfung der korrekten Klonierung

#### Transformation chemisch-kompetenter *E. coli*-Zellen

Zur Überprüfung der korrekten Klonierung wurde 0,15  $\mu\text{l}$  des Ligationsansatzes in kompetente *E. coli* XL10 Gold-Zellen transformiert. Analog wurde mit den Zellen *E. coli* Rosetta (DE3) zur Expression verfahren. Dazu wurde der Transformationsansatz 20 min auf Eis inkubiert und im Anschluss daran für 45 s bei 42 °C ein Hitzeschock durchgeführt. Es folgte eine Inkubation für 2 min auf Eis. Daraufhin wurde der Ansatz mit 200  $\mu\text{l}$  LB-Medium (Raumtemperatur) versehen und für 1 h bei 37 °C und 500 rpm schüttelnd im Thermomixer (Eppendorf) inkubiert. 100  $\mu\text{l}$  wurden auf einer mit 50  $\mu\text{g}/\text{ml}$  versetzten LB-Agar-Platte mit Hilfe von sterilen Glasperlen ausplattiert und über Nacht bei 37 °C inkubiert. Am nächsten Tag auf der Platte sichtbare Kolonien wurden für eine Plasmid-Minipräparation angezogen und die DNA nach Anleitung des NucleoSpin Kits (Macherey-Nagel) isoliert.

#### Anzucht von *E. coli*, Plasmid-Minipräparation und Restriktion

Ausgehend von den auf Agarplatten angezogenen *E. coli*-Kulturen wurde zur Plasmid-Minipräparation eine Flüssigkultur angelegt. Dazu wurde eine Einzelkolonie mit einer Pipettenspitze abgeimpft und in ein steriles Reagenzglas, gefüllt mit je 8 ml LB-Medium und 50  $\mu\text{g}/\text{ml}$  Kanamycin, gegeben und über Nacht bei 37 °C und 200 rpm schüttelnd inkubiert. Am nächsten Morgen wurde davon eine Plasmid-Minipräparation nach Anleitung des Plasmid-DNA-Purification-Kits (Macherey-Nagel) durchgeführt und 0,1  $\mu\text{g}$  der extrahierten DNA mit *ApaI* und *XhoI* restringiert. Dazu wurden 0,1  $\mu\text{g}$  DNA mit 2  $\mu\text{l}$  *Fast Digestion Buffer* versetzt und 0,5  $\mu\text{l}$  *XhoI* ((*Fast-Digestion*-Enzym, Fermentas) sowie 1  $\mu\text{l}$  *ApaI* (*Fast-Digestion*-Enzym, Fermentas) hinzugefügt, der Ansatz auf 20  $\mu\text{l}$  mit Aqua bidest aufgefüllt, 10 min bei 37 °C inkubiert und im Anschluss auf ein 1,5 %iges Agarosegel aufgetragen.

80 ng/ $\mu\text{l}$  potentiell positiver Klone wurden zum Sequenzieren der Firma GATC geschickt. Die zur Sequenzierung verwendeten Primer waren pET\_RP und pGEX5\_FP.

### 2.3.3. Expression von *mut8* in *E. coli* Rosetta (DE3)

Die Anzucht erfolgte hier ausgehend von den transformierten *E. coli* Rosetta (DE3) (siehe Kap. 2.3.2). Dazu wurden 30 ml LB-Medium, mit 50  $\mu\text{g/ml}$  Kanamycin versetzt, in einen 250 ml Kolben gegeben und mit einer Einzelkolonie des Transformanten angeimpft. Der Ansatz wurde über Nacht bei 37 °C und 180 rpm schüttelnd inkubiert. Am nächsten Morgen wurde 1 l LB-Medium (im 5 l-Kolben) mit 50  $\mu\text{g/ml}$  Kanamycin versetzt und mit der Vorkultur angeimpft. Dieser Ansatz wurde etwa 2-3 h bei 37 °C und 180 rpm schüttelnd inkubiert, bis eine  $\text{OD}_{600}$  von 0,8 erreicht wurde. Im Anschluss daran wurden die Zellen für 15 min bei 5000 rpm und 21 °C abzentrifugiert (JLA 9.1000) und das Pellet vorsichtig in etwa 50 ml M9-Medium resuspendiert. Das resuspendierte Pellet wurde zu den restlichen 4 l M9-Medium gegeben und auf 4 Kolben aufgeteilt (je 1 l/Kolben). Diese Hauptkultur wurde bei 37 °C und 180 rpm schüttelnd bis zu einer  $\text{OD}_{600}$  von 1,5 inkubiert und mit 0,4 mM IPTG die Expression von *mut8* induziert. Die Expression erfolgte über 5 h, wobei die Kultur bei 28 °C und 180 rpm schüttelnd inkubiert wurde. Die angezogenen Zellen wurden nach der Expression von *mut8* für 15 min bei 5000 rpm und 4 °C abzentrifugiert (JLA 9.1000), das Pellet in ein 50 ml-Falcon-Röhrchen überführt und bis zum Zellaufschluss bei -20 °C (siehe Kap. 2.4.1) gelagert.

## 2.4. Biochemische Methoden

### 2.4.1. Zellaufschluss

Die nach der Expression von *mut8* bei -20 °C gelagerten Zellen (siehe Kap. 2.3.3) wurden bei Raumtemperatur (RT) aufgetaut, mit 5 mg Lysozym versetzt und mit Waschpuffer auf 50 ml aufgefüllt. Der Ansatz wurde rollend bei RT für 1 h inkubiert und die Zellen anschließend im Microfluidizer mechanisch zerstört. Dafür wurden sie 3x hintereinander durch die Zellaufschlussanlage gepumpt, bis die erhaltene Zellsuspension klarer wurde. Im Anschluss daran wurden die so aufgeschlossenen Zellen bei 30000 rpm und 4 °C für 1 h zentrifugiert (Ti45-Rotor, Beckman Ultrazentrifuge). Der Überstand wurde steril filtriert (Millipore Express, 0,22  $\mu\text{m}$ -Membran), um ihn von Schwebeteilchen zu befreien, und stand damit zur Proteinreinigung zur Verfügung.

### 2.4.2. Proteinreinigung

#### GSH (Glutathion)-Affinitätschromatographie

Der erste Reinigungsschritt erfolgte mittels zweier hintereinander gekoppelter GSH (Glutathion)-Säulen (GSTrap HP, GE-Healthcare) an der ÄktaFPLC-Anlage. Dafür wurde der filtrierte Überstand nach dem Ultrazentrifugationsschritt (siehe oben) über einen 150 ml-Glasloop auf die Säule gegeben, die Chromatographie nach Tab. 2.3 durchgeführt und die Elution anhand der Extinktion bei 280 nm verfolgt. Von den Peakfraktionen wurde je 5  $\mu\text{l}$  abgenommen und für ein nachfolgendes SDS-Gel aufbewahrt (siehe Kap. 2.4.4). Die Peakfraktionen wurden vereinigt, auf 500  $\mu\text{l}$  aufkon-



**Tabelle 2.3.:** Programm der Chromatographie von mut8, die verwendete GSH-Säule hat ein Volumen von 1 ml. Zwei dieser Säulen wurden direkt hintereinander geschaltet, somit ergibt sich ein Gesamt-Säulenvolumen von 2 ml.

	Puffer	Volumen (ml)	Flussrate (ml/min)
Äquilibrieren	Waschpuffer	24	1
Auftragen	-	250	1
Waschen I	Hochsalzpuffer	40	0,75
	Waschpuffer	40	1
Eluieren	Elutionspuffer	40	0,5
Regenerieren	Guanidiniumchlorid (6 M)	24	0,75
	PBS (1x)	30	1
	Aqua bidest	60	1
Lagern	Ethanol (20 %)	24	0,5

zentriert (Millipore, 10.000 MCWO) und ein Aliquot zur Proteinkonzentrationsbestimmung (siehe Kap. 2.4.3) entnommen. Die Proteinkonzentration wurde nach der Bradford-Methode bestimmt.

Zu dem so gereinigten GST\_mut8 wurden 2,3  $\mu\text{g}/(\text{mg GST\_mut8})$  PreScission-Protease gegeben und der Ansatz bei 4 °C über Nacht inkubiert, um den GST-tag von mut8 zu trennen.

### Gelfiltrationschromatographie

Nach der GSH-Affinitätschromatographie wurde das GST-mut8-Gemisch über eine Gelfiltrationschromatographie getrennt. Verwendet wurde dafür eine Sephadex 75pg (16/60) (GE-Healthcare) mit einem Säulenvolumen von 125 ml. Die Reinigung erfolgte nach Tabelle 2.4, die Elution wurde anhand der Extinktion bei 280 nm verfolgt.

**Tabelle 2.4.:** Programm der Gelfiltrationschromatographie von mut8, die verwendete Säule HiLoad 16/60 Superdex 75pg (GE-Healthcare) hat ein Volumen von 125 ml.

	Puffer	Volumen (ml)	Flussrate (ml/min)
Äquilibrieren	Gelfiltrationspuffer	150	1
Auftragen	-	1	0,5
Eluieren	Gelfiltrationspuffer	200	0,75
Waschen	Aqua bidest	150	1
Lagern	Ethanol (20 %)	150	0,5

Je 10  $\mu\text{l}$  der proteinhaltigen Fraktionen wurden für ein folgendes SDS-Gel aufbewahrt. Die Peakfraktionen wurden vereinigt und auf 500  $\mu\text{l}$  aufkonzentriert (Sartorius, 2.000 MCWO). Während des Einkonzentrierens wurde die Probe durch dreimaliges Auffüllen des Centricons mit 50 mM KPi-Puffer, pH 6,5, umgepuffert, um mut8 in

einem NMR-adäquaten Puffer vorliegen zu haben. Nach dem Einkonzentrieren auf 540  $\mu\text{l}$  wurde ein Aliquot zur Proteinkonzentrationsbestimmung (siehe Kap. 2.4.3) entnommen. Die Probe stand nun bereit für die NMR-Messungen (siehe Kap. 2.5.1).

### 2.4.3. Proteinkonzentrationsbestimmung

Die Proteinkonzentrationsbestimmung erfolgte nach der Reinigung über die GSH-Säule über den Bradford-Test. Dazu wurden 999  $\mu\text{l}$  Bradford-Reagens in eine Halbmikroküvette gegeben, 1  $\mu\text{l}$  Proteinprobe hinzugegeben, der Ansatz vorsichtig geschwenkt und 3 min lang bei Raumtemperatur inkubiert. Die Proteinkonzentration wurde bei 295 nm im Biophotometer (Eppendorf) gegen einen BSA-Standard gemessen.

Nach der abschließenden Reinigung über Gelfiltrationschromatographie wurde der Proteingehalt nach Lambert-Beer mit  $\epsilon=8480 \text{ M}^{-1}\text{cm}^{-1}$  bestimmt. Dafür wurden den gesammelten und aufkonzentrierten Fraktionen 10  $\mu\text{l}$  entnommen, zu 60  $\mu\text{l}$  Aqua bidest einer UV-Küvette (Halbmikro Einmalküvette, ratiolab) gegeben und die Extinktion E bei 280 nm gemessen. Unter Einbeziehen des Verdünnungsfaktors ergibt sich nach Gleichung 2.1 für die Proteinkonzentration c (mit d=1 cm):

$$c = E_{280}/d \cdot \epsilon \quad (2.1)$$

### 2.4.4. SDS-PAGE

Für die SDS-Polyacrylamidgelelektrophorese wurden 12,5 %ige Gele, die sich folgendermaßen zusammensetzen (siehe Tab. 2.5), eingesetzt. Die Angaben beziehen sich dabei immer auf je vier Trenngele und fünf Sammelgele für die Gelapparatur Mini Protean 3 Cell von Bio-Rad. Als Größenmarker wurden der 10-250 kDa Protein-Marker (New England Biolabs) und der Mark12 *Unstained Standard* (Invitrogen) verwendet (siehe Abb. 2.1, Mitte, rechts).

**Tabelle 2.5.:** Zusammensetzung unterschiedlich konzentrierter Polyacrylamidgele.

	<b>Trenngel (4x)</b> <b>15 %</b>	<b>Trenngel (4x)</b> <b>12,5 %</b>	<b>Sammelgel (5x)</b>
Trenngelpuffer	4,4 ml	4,4 ml	-
Sammelgelpuffer	-	-	500 $\mu\text{l}$
Aqua bidest.	5,2 ml	6,9 ml	3,5 ml
10 % SDS	200 $\mu\text{l}$	200 $\mu\text{l}$	50 $\mu\text{l}$
AA/Bis	10 ml	8,3 ml	850 $\mu\text{l}$
Temed	8 $\mu\text{l}$	8 $\mu\text{l}$	5 $\mu\text{l}$
APS	200 $\mu\text{l}$	200 $\mu\text{l}$	50 $\mu\text{l}$

Nach der SDS-Gelelektrophorese werden die Gele vorsichtig aus der Kammer gelöst, das Sammelgel abgetrennt und das übrige Trenngel für 15 min schwenkend in einer

Coomassie-Lösung gefärbt. Anschließend erfolgt eine 30-minütige Entfärbung mit Entfärbelösung, ein Wechsel dieser und ein erneutes Entfärben über Nacht, so dass die blau gefärbten Proteinbanden deutlich zu erkennen sind.

## 2.5. NMR-Spektroskopie

### 2.5.1. Probenvorbereitung

Für die Aufnahme des 1D- sowie aller homonuklearen Spektren und dem  $^1\text{H}$ - $^{15}\text{N}$ -SOFAST-HMQC (im Folgenden nur als SOFAST-HMQC bezeichnet) wurde das synthetische mut8 in einer Endkonzentration von 1 mM in 600  $\mu\text{l}$  KPi-Puffer, pH 6,5, 10 % (v/v)  $\text{D}_2\text{O}$ , 0,02 % (w/v)  $\text{NaN}_3$  und 50  $\mu\text{M}$  DSS (2,2-Dimethyl-2-silapentanon-5-sulfonat) gelöst und in ein NMR-Röhrchen gefüllt (5 mm-ThinWall-Probenröhrchen, Ampol). Die Probe wurde nach vollständiger Aufnahme aller zuvor genannten Spektren gefriergetrocknet (siehe Kap. 2.5.2) und in 100 %  $\text{D}_2\text{O}$  (600  $\mu\text{l}$ ) gelöst.

Für die Aufnahme des heteronuklearen  $^1\text{H}$ - $^{15}\text{N}$ -HSQCNOE (im Folgenden nur als HSQCNOE bezeichnet) wurde das 540  $\mu\text{l}$ -Aliquot des  $^{15}\text{N}$ -markierten mut8 ( $\approx 200 \mu\text{M}$ ) (siehe Kap. 2.4.2) mit 60  $\mu\text{l}$   $\text{D}_2\text{O}$  versetzt und in ein NMR-Röhrchen überführt (5 mm-ThinWall-Probenröhrchen, Ampol).

### 2.5.2. Gefriertrocknung von mut8

Das Peptid mut8 wurde zur Messung in 100 %  $\text{D}_2\text{O}$  in ein Eppendorfgefäß überführt und in flüssigem Stickstoff schockgefroren. Die vollständige Trocknung erfolgte an der Gefriertrocknungsanlage (Christ) über 48 h bei einer Temperatur von  $-62^\circ\text{C}$ . Das gefriergetrocknete Peptid wurde für eine weitere NMR-Messung in 100 %  $\text{D}_2\text{O}$  gelöst (siehe Kap. 2.5.1).

### 2.5.3. Messbedingungen

Alle Messungen wurden am Bruker Ultrashield 700-NMR (Bruker BioSpin, Fällanden) bei einer Protonenfrequenz von 700,22 MHz durchgeführt. Verwendet wurde dabei ein inverser Tripleresonanz-Cryo-Probenkopf (700 MHz) mit aktiv abgeschirmten Gradientenspulen in allen drei Raumrichtungen. Die Temperatur betrug bei allen Messungen 298 K und wurde über eine BVT3000-Kontrolleinheit (Bruker BioSpin, Fällanden) gesteuert. Die Spektren wurden alle im phasensensitiven Modus mit Quadraturdetektion in beiden Zeitdimensionen unter Verwendung von States-TPPI (*States-time-proportional-phase-incrementation*) [81] bzw. Echo-Antiecho aufgenommen. Die Wasserresonanz wurde bei allen Experimenten bis auf dem 1D- $^1\text{H}$ - und dem  $^1\text{H}$ - $^1\text{H}$ -COSY-Experiment (im Folgenden nur als COSY bezeichnet) über die 3-9-19-WATERGATE-Pulsfolge [82] unterdrückt. Die Mischzeiten betrugen für das  $^1\text{H}$ - $^1\text{H}$ -NOESY-Experiment (im Folgenden nur als NOESY bezeichnet) 200 ms und für das  $^1\text{H}$ - $^1\text{H}$ -TOCSY-Experiment (im Folgenden nur als TOCSY bezeichnet) 80 ms. Der

**Tabelle 2.6.:** NMR-Experimente und Parameter. 1: Zahl der gemessenen *scans*, 2: spektrale Breite in Hz, 3: Zahl der realen Punkte in der jeweiligen Dimension.

Experiment	F1			F2		Referenz
(Pulsprogramm)	NS <sup>1</sup>	SW <sup>2</sup> (Hz)	TD <sup>3</sup>	SW <sup>2</sup> (Hz)	TD <sup>3</sup>	
1 mM mut8 in 50 mM KPi, pH6,5						
<sup>1</sup> H (zgpr)	128	8417,51	16834	-	-	
COSY (cosygppr)	64	10504,19	512	10504,20	2048	[83]
TOCSY (dipsi2etgpsi19)	64	11203,59	512	11160,71	2048	[84, 76, 74]
NOESY (noesygppl19)	64	10504,19	512	10504,20	2048	[85]
SOFAS-HMQC (sfhmqcf3gppl)	2000	2412,67	256	9803,92	1024	[86]
200 μM <sup>15</sup> N-mut8 in 50 mM KPi, pH6,5						
HSQCNOE (hsqcnoef3gpsi)	16	2838,43	1024	11261,26	2048	[87, 88]

90° Hochleistungs-Puls (p1) betrug 11,5  $\mu$ s. Der 120°-*shaped pulse* (sp23) betrug bei Durchführung des SOFAST-HMQCs 33,7 dB, der 180°-*shaped pulse* (sp24) betrug 25,4 dB. Alle verwendeten Pulsprogramme wurden der Bruker Standardbibliothek entnommen. Das der Probe zugesetzte D<sub>2</sub>O diente als interner Feldfrequenz-Lock. Als externer Standard diente das DSS (2,2-Dimethyl-2-silapentan-5-sulfonat), gegen welches die ppm-Werte der Protonenfrequenzen geeicht wurden. Tabelle 2.6 listet alle durchgeführten NMR-Experimente, die dazu gehörigen Pulsprogramme und wichtigsten Parameter auf.

Eine in 50 mM KPi, pH 6,5 gelöste Peptidprobe wurde lyophilisiert (siehe Kap. 2.5.2) und zum Nachweis von langsam austauschenden Amidprotonen in 100 % D<sub>2</sub>O gelöst. Es wurde ein NOESY (siehe Tabelle 2.6) aufgenommen.

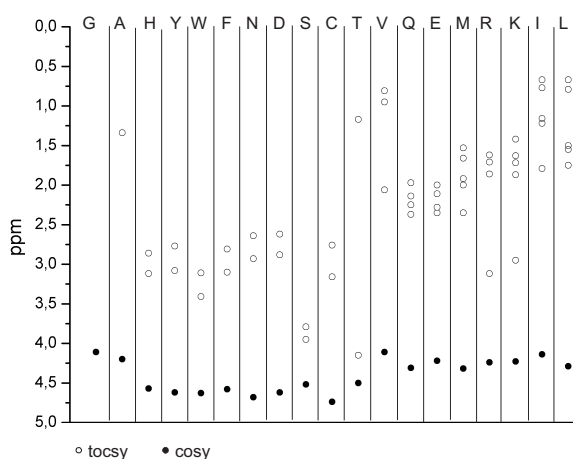
### 2.5.4. Prozessierung der Spektren

Das Programm Topspin 2.0 (Bruker BioSpin, Rheinstetten) wurde sowohl zur Aufnahme als auch zur Prozessierung der Spektren verwendet. Alle FIDs (*free induction decays*) wurden vor der Fourier-Transformation (FT) in F1 auf eine Menge 2048 Datenpunkte aufgefüllt (*zero-filling*), um die digitale Auflösung der zweidimensionalen Spektren zu verbessern.  $\pi/4$ - bis  $\pi/2$ -verschobene Sinus- oder Sinusquadratfunktionen wurden als Filterfunktion verwendet [89]. Die Basislinienkorrektur erfolgte standardmäßig über den Topspin-Befehl *abs* durch Subtraktion einer Polynomfunktion

fünften Grades in beiden Dimensionen F1 und F2. Die gleichzeitig alternierend aufgenommenen HSQCNOE-Spektren wurden über den Befehl *split 2* in zwei Spektren (Spektrum mit Vorsättigung ( $I_{sat}$ ) bzw. Spektrum ohne Vorsättigung ( $I_{eq}$ )) geteilt.

### 2.5.5. Auswertung der Spektren

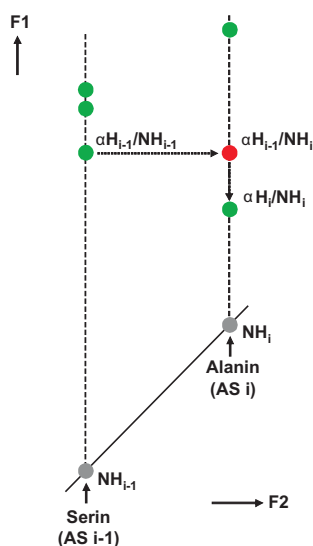
#### Sequentielle Zuordnung der Protonenresonanzen



**Abbildung 2.4.:** Typische Signalmuster der einzelnen Aminosäuren. Dargestellt sind schematisch die Frequenzen (ppm) der Kopplungen zwischen dem  $H_N$  und den  $H_{\alpha, \beta, \gamma, \delta, \epsilon}$  der 19 Aminosäuren innerhalb eines Pentapeptids (GXAGG). Prolin ist aufgrund des fehlenden  $H_N$  nicht mit aufgelistet. Angelehnt an [90].

Nach der Spektrenprozessierung wurden das COSY-, TOCSY- und NOESY-Spektrum in das Programm AUREMOL [78] eingelesen und überlagert. Mit Hilfe von Tabelle 1 aus Wishart *et al.*, 1995 [90] wurde eine Zuordnungsübersicht der einzelnen Aminosäuren erstellt und damit die Signale der Spektren zugeordnet (siehe Abb. 2.4).

Im COSY-Spektrum sind vicinale Kopplungen, also Kopplungen über maximal drei chemischen Bindungen zwischen zwei Atomen, zu sehen. Im TOCSY wird der Magnetisierungstransfer über alle durch skalare Kopplungen verbundenen Protonen eines Spinsystems (= Aminosäure) detektiert, sodass hierdurch eine eindeutige Identifizierung der jeweiligen Aminosäure möglich wird. NOE-Kreuzresonanzen entstehen über dipolare Kopplungen im Raum bis zu einem Abstand von  $\geq 5 \text{ \AA}$ , sodass im NOESY-Spektrum neben den Resonanzen des eigenen Spinsystems auch die von sequentiell oder strukturell angrenzenden Protonen sichtbar sind. In Kombination mit der Primärsequenz des Peptids können mit Hilfe des *Chaintracing*-Verfahrens nach Wüthrich, 1986 [91] über die  $H_{\alpha}$ - $H_N$ -Kreuzresonanzen alle Aminosäuren sequentiell zugeordnet werden. Das Prinzip des Verfahrens ist in Abbildung 2.5 dargestellt.



**Abbildung 2.5.:** Prinzip des *Chaintracing*-Verfahrens nach Wüthrich, 1986 [91]. Zu sehen sind in grün die theoretisch zu erwartenden TOCSY-Signale für ein Alanin (Aminosäure  $i$ ) und dem N-terminal direkt benachbarten Serin (Aminosäure  $i-1$ ) (siehe auch Abb. 2.4). In rot ist das NOESY-Signal dargestellt, welches Aminosäure  $i-1$  (Serin) mit Aminosäure  $i$  (Alanin) verbindet.

### Zuordnung der heteronuklearen Spektren

Die Zuordnung der heteronuklearen Spektren erfolgte mit Hilfe des der  $H_\alpha$ -Werte des bereits zugeordneten TOCSY-Spektrums und nach Wishart *et al.*, 1995 [90].

### Ableitung der Distanzbeschränkungen

Mit Hilfe des Programms AUREMOL [78] wurden nach der vollständigen Zuordnung aller Signale im NOESY-Spektrum die Volumina dieser berechnet. Die Integration erfolgt dabei unter Berücksichtigung eines aus der durchschnittlichen Halbwertsbreite der Signale (*line width at half height*) berechneten Standardfehlers für Volumen und Form der Resonanzsignale. Kalibriert werden die NOE-Signalintensitäten gegen den Abstand zwischen den  $\delta$ - und  $\epsilon$ -Protonen einer im Protein enthaltenen aromatischen Aminosäure [91] und dann eine Liste der NOEs mit Abständen in Å erstellt. Die Werte der so erstellten Liste gingen als Randbedingungen in die anschließende Strukturberechnung (siehe Kap. 2.6) mit ein.

### Vorhersage der Diederwinkel über TALOS

TALOS erstellt über das Zugreifen auf eine Datenbank eine empirische Vorhersage von  $\phi$ - und  $\psi$ -Winkeln für das zu untersuchende Protein anhand von zugeordneten  $C_\alpha$ -,  $C_\beta$ -, CO- und N-Resonanzen. In dieser Arbeit wurden nur die  $H_\alpha$ -Resonanzen berücksichtigt, da nur diese dem homonuklearen Datensatz entnommen werden konnten. TALOS nutzt das Wissen, dass die chemischen Verschiebungen in Zusammenhang mit

der Sekundärstruktur des jeweiligen Proteins stehen. Demnach sollten Aminosäuren bei ähnlichen Verschiebungen im Spektrum auch ähnliche Torsionswinkel aufweisen. Ziel ist es, die chemischen Verschiebungen und Sequenzinformationen zu verwenden, um eine quantitative Vorhersage für die Torsionswinkel  $\phi$  und  $\psi$  zu erstellen. Die chemische Verschiebungen des Proteins werden in Triplet-Schritten mit den Verschiebungen, abgelegt in einer Datenbank, verglichen. Jede Aminosäure erhält dann den Durchschnittswert der jeweils 10 bestgeeigneten Vorhersagen. Diese Vorhersagen können dann vom TALOS-Benutzer als „gut“, „mittel“ oder „schlecht“ klassifiziert werden. Die Diederwinkel-Vorhersagen flossen als zusätzliche Randbedingungen ebenfalls mit in die Strukturberechnung (siehe Kap. 2.6) ein.

### Analyse der Sekundärstruktur nach Wüthrich

Das Vorhandensein bestimmter NOEs innerhalb des NOESY-Spektrums erlaubt es, ein strukturelles Element in einem Peptid oder Protein als  $\alpha$ -Helix oder  $\beta$ -Faltblatt zu deklarieren [92]. Sequentiell aufeinanderfolgende NOEs zwischen Amidprotonen der Aminosäure  $i$  und Aminosäure  $i+2$  ( $d_{NN}(i, i+2)$ ) - sowie von  $d_{\alpha N}(i, i+3)$ ,  $d_{\alpha N}(i, i+4)$  und  $d_{\alpha\beta}(i, i+3)$  NOEs sind charakteristisch für  $\alpha$ -helikale Bereiche. Aus der Liste aller zugeordneten Resonanzen wurden besagte NOEs extrahiert und die Intensitäten in „schwach“, „mittel“ und „stark“ eingestuft. Je nach Vorhandensein dieser NOEs wurde ein Bereich als  $\alpha$ -Helix festgelegt oder bei Fehlen von NOEs als Zufallsknäuel deklariert.

### Analyse der Sekundärstruktur nach Wishart

Die für die chemische Verschiebung der  $^1\text{H}_\alpha$  ermittelten Werte wurden für eine Sekundärstrukturanalyse verwendet und der chemische Verschiebungsindex (CSI) ermittelt. Die hier beschriebene Analysemethode erlaubt es, anhand der chemischen Verschiebungen der  $\alpha$ -Protonen eines Peptides oder Proteins relativ zu einem Standard wie 2,2-Dimethyl-2-silapentan-5-sulfonsäure (DSS),  $\alpha$ -Helices oder  $\beta$ -Faltblätter vorherzusagen. Verglichen werden bei dieser Methode die Resonanzfrequenzen der  $\alpha$ -Protonen der Aminosäuren mit den entsprechenden Frequenzen eines Pentapeptids GXAGG. Befinden sich die Aminosäuren in einer helikalen Anordnung, so sind die  $\alpha$ -Protonen gegenüber den Frequenzen der Aminosäuren des Pentapeptids hochfeldverschoben. Befinden sich die Aminosäuren in einer  $\beta$ -Faltblatt-Anordnung, so sind die Resonanzfrequenzen tieffeldverschoben [93]. Bei einer Abweichung des gemessenen Wertes vom Referenzwert um mehr als  $\pm 0,1$  ppm erhält das entsprechende  $\text{H}_\alpha$  den Wert  $\pm 1$ . Bei einer Abweichung  $< 0,1$  ppm wird dem entsprechenden  $\text{H}_\alpha$  der Wert 0 zugeordnet. Drei oder mehr aufeinander folgende +1-Werte, die nicht durch eine -1 unterbrochen werden, sprechen für einen  $\beta$ -Faltblatt Bereich, folgen vier oder mehr -1-Werte aufeinander, die nicht durch eine +1 unterbrochen werden, so liegt in dem Bereich eine  $\alpha$ -Helix vor. Der CSI wird graphisch gegen die Aminosäureposition aufgetragen, so können die einzelnen Sekundärstrukturelemente identifiziert werden.

## 2.6. Strukturaufklärung und Dynamik

Eine computerbasierte Strukturaufklärung von Peptiden und Proteinen benötigt neben den NMR-Randbedingungen (NOE-Konnektivitäten, -Intensitäten, s. Kap. 2.5.5) und Diederwinkeln (siehe Kap. 2.5.5) zusätzliche Parameter [94]. So werden zusätzlich empirische Konstanten für Bindungslängen, Geometrie und nichtkovalente Wechselwirkungen mit einbezogen.

### 2.6.1. Das Xplor-NIH-Kraftfeld

Zur Strukturaufklärung wurde in dieser Arbeit das Programm Xplor-NIH verwendet [76, 77]. Xplor-NIH stellt ein allgemeines Paket zur Strukturberechnung von Biomolekülen aus experimentellen NMR-Daten kombiniert mit bekannten geometrischen Daten dar und wurde aus dem MD- und Minimierungsprogramm CHARMM entwickelt. Eine Struktur wird darüber bestimmt, dass das Minimum einer Zielfunktion gesucht wird, welche aus Termen für die experimentellen NMR-Beschränkungen, kovalenten Geometrien und nicht-Bindungs-Kontakten besteht. Die hier verwendete Zielfunktion beinhaltet folgende energetische Terme (siehe Gleichung 2.2):

$$E_{pot} = E_{bonds} + E_{angle} + E_{dih} + E_{impr} + E_{vdW} + E_{exp} \quad (2.2)$$

Für die einzelnen Energien gilt Folgendes:

Gesamtterm der potentiellen Energie	$E_{pot}$
Bindungspotential	$E_{bonds} = \sum k_b \cdot (r - r_0)^2$
Winkelpotential	$E_{angle} = \sum k_\theta \cdot (\theta - \theta_0)^2$
Diederwinkelpotential	$E_{dih} = \sum k_\phi \cdot (1 + \cos(n\phi + \phi_0))$
Improperwinkel-Potential	$E_{impr} = \sum k_\omega \cdot (\omega - \omega_0)^2$
Van der Waals-Potential	$E_{vdW} = \sum \left( \frac{A}{r_{ij}^{12}} - \frac{B}{r_{ij}^6} + \frac{q_i q_j}{D r_{ij}} \right)$
Experimentelle Randbedingungen	$E_{exp}$

Der experimentelle Energieterm ergibt sich aus  $E_{exp} = E_{NOE} + E_{dih}$ , wobei  $E_{dih}$  hier die über TALOS bestimmten Diederwinkel-Konformationen bezeichnet. Für das NOE-Potential gilt dann Folgendes:

$$E_{NOE} = \sum k_{NOE} \cdot (r - r_i^0)^2 \quad (2.3)$$

Die einzelnen Buchstaben bedeuten:  $k_x$ , jeweilige Kraftkonstante; b, aktuelle Bindungslänge;  $\theta$ , aktueller Bindungswinkel;  $\phi$ , dihedraler Winkel mit  $n=1,2,3,4,6$ ;  $\omega$ , Improperwinkel;  $r_{ij}$ , aktueller Abstand der Kerne i und j voneinander;  $b_0$ ,  $\theta_0$ ,  $r_{ij}^0$ , die jeweiligen Referenzwerte; A,B, Lennard-Jones-Terme; qx, Ladung des Kerns x; D, modifizierte Dielektrizitätskonstante.

### 2.6.2. Erstellen der Koordinatendateien

Um Proteinstrukturen zu berechnen, wird durch Minimieren der Zielfunktion 2.2 ein globales Minimum gesucht. Die Läufe basierten auf einer Startstruktur mit ausgestreck-



ter Konformation, um eventuell auftretende konformationelle Nebenminima, welche strukturelle Artefakte darstellen, zu vermeiden. Das Protokoll *generate.inp* (siehe Anhang A) erzeugte dafür eine .psf-Datei. In dieser sind die wichtigen Parameter für die Minimierungsrechnungen im Kraftfeld enthalten. Dazu zählen die Peptidsequenz, die sequentielle Anordnung der Atome in den Aminosäuren mit den jeweiligen Winkeln, Atomtypen-Zuordnung, Atommassen, Ladungen und Improperwinkel. Über das Skript *generate\_template.inp* wurde eine .pdb-Datei erzeugt (siehe Anhang A).

### 2.6.3. *Simulated Annealing*-Läufe

*Simulated Annealing-Läufe* simulieren Abkühlungen, um hochkomplexe Funktionen zu minimieren. Das System, welches anfangs auf sehr hohe Temperaturen aufgeheizt wurde, wird langsam abgekühlt. Die sehr hohe Anfangsenergie des Systems sorgt dafür, dass lokale Minima überwunden werden, während sich durch die kontrollierte Abkühlungsphase sehr wahrscheinlich Lösungen am globalen Minimum ergeben. Diese Lösungen am globalen Minimum der Zielfunktion entsprächen dann einem Ensemble an möglichen Strukturen des Proteins bzw. Peptids. Das hierfür verwendete Protokoll nennt sich *sa\_mut8.inp* (siehe Anhang A). Dort wird eine Energieminimierung (*Conjugate Gradient* (CG)-Algorithmus) durchgeführt [95], im Anschluss daran folgt ein kurzer Hochtemperatur-Molekulardynamiklauf bei 1000 K. Abschließend erfolgt eine Abkühlphase, in der das System in 3000 Schritten auf 100 K gebracht wird. Dabei erhöht sich die Wichtung für die dihedralen Terme um das 200fache. Über eine letzte CG-Minimierung in 200 Schritten wird eine Optimierung der Struktur erreicht. Insgesamt wurde so eine Schar von 100 Strukturen berechnet.

### 2.6.4. Verfeinerung der Struktur

Nach Durchführung des *sa\_mut8.inp*-Protokolls wurde eine Verfeinerung der Strukturen über das *ref\_190100.inp*-Protokoll durchgeführt (siehe Anhang A). Dies soll die erhaltenen Winkel- und Konformationsgeometrien für alle 100 berechneten Strukturen (Kap. 2.6.3) optimieren. Die Abkühlphase wird in diesem Protokoll mit 4000 Schritten (statt 3000 Schritten wie bei *sa\_mut8.inp*) durchgeführt und vermindert somit, dass thermodynamisch ungünstige Konformationen entstehen. Dabei werden während der anfänglichen Hochtemperaturphase die experimentellen NOE-Randbedingungen und Bindungslängen mehr gewichtet und während der Abkühlphase die Skalierung der hydrophoben Wechselwirkungen verringert. Nach der Verfeinerung wurden die 10 besten Konformere aus dem 100er-Ensemble ausgewählt. Als Auswahlkriterien galten minimale Energie und möglichst keine Verletzungen der experimentellen Randbedingungen.

### 2.6.5. Analyse der Ensembles

Zur Analyse der 10 besten Strukturen des Ensembles wurden diese in RamPage [1] eingelesen und so Ramachandranplots [96] erstellt. Darüber wurde der prozentuale

Anteil an Aminosäuren, die sich in favorisierten, erlaubten oder nicht erlaubten Bereichen aufhalten, berechnet. Ebenfalls wurden alle 10 Strukturen in MoleMan [74] eingelesen, welches eine prozentuale Bestimmung der im Peptid enthaltenen Sekundärstrukturelemente ermöglicht.

### 2.6.6. Dynamik-Berechnungen

Um mut8 hinsichtlich seiner Dynamik charakterisieren zu können, wurde das aufgenommene HSQCNOE-Spektrum (siehe Kap. 2.5.3) ausgewertet. Über das Pulsprogramm `hsqnoef3gpsi` [87, 88] wurden gleichzeitig alternierend HSQCNOE-Spektren aufgenommen, welche über den Topspin-Befehl `split 2` in zwei Spektren geteilt wurden. Ein Spektrum besteht lediglich aus positiven Signalen ( $I_{eq}$ ), das zweite Spektrum aus negativen und positiven Signalen ( $I_{sat}$ ).

Die Signalintensität des erzeugten NOEs hängt von verschiedenen Bewegungen ab, die der jeweilige  $H_N$ -Bindungsvektor durchführt. Neben einem Term, der sich aus der Gesamtrationsbewegung des Peptids als starre Kugel (*tumbling*) ergibt, wird der NOE von den lokalen Fluktioationsbewegungen (Librationsbewegungen, Kettentorsionen und freie Translationsbewegungen von Schleifen oder Termini) beeinflusst. Der resultierende NOE spiegelt daher alle Bewegungsmodi wider, die die dipolare Kopplung auf der ns-Skala beeinflussen können. Diese komplexe Einbeziehung aller dieser Modi führt zu positiven und negativen Resonanzen im hetNOE-Spektrum. Daher lässt sich, nach Interpretation und Auswertung der Spektren, eine Gesamtbewegung bestehend aus den Fluktuationen der  $H_N$ s, der Bewegungen der Sekundärstrukturelemente wie  $\alpha$ -Helices und des ganzen Proteins bzw. Peptids berechnen. Dieser hetNOE (*heteronuclear*NOE) ( $\eta$ ) wird berechnet, indem der Quotient aus  $I_{sat}$  und  $I_{eq}$  gebildet wird (siehe Gleichung 2.4).

$$\eta = I_{sat}/I_{eq} \quad (2.4)$$

Zur Bestimmung der Signalintensitäten werden beide Spektren in das Programm Sparky eingelesen [79] und über eine Gauß-Funktion die Volumina der einzelnen Resonanzen ausgelesen.

### 3. Ergebnisse

Proteine und Peptide können *in silico* mit Hilfe von evolutionären Algorithmen optimiert werden. Dabei ist es sinnvoll, während der Optimierung mehrere Parameter gleichzeitig verbessern zu können. Ein erzeugter Vorteil einer einzigen Optimierung – wie z. B. eine Erhöhung der strukturellen Stabilität des Peptids – könnte sonst den Verlust einer anderen Eigenschaft nach sich ziehen, so kann ein vorher gut in wässriger Umgebung lösliches Peptid nach der Optimierung schlechter löslich sein, obwohl es gleichzeitig strukturell stabiler geworden ist. Um zu gewährleisten, dass mehrere Parameter gleichzeitig optimiert werden können, muss ein entsprechender Algorithmus daher multiobjektiv sein. Der von Hohm *et al.*, 2005 entwickelte Algorithmus EMOS arbeitet multiobjektiv und wurde bereits am Modellpeptid Villin Headpiece 36 (HP36) getestet [27, 9].

Das von Gronwald *et al.* (2008) vorgestellte Resultat eines Optimierungsprozesses über EMOS ist ein Cluster an möglichen Sequenzen. Diese unterscheiden sich gegenüber der Wildtypsequenz (HP36, pdb-ID: 1vii) durch Austausch einer willkürlichen Anzahl an Aminosäuren und wurden auf mindestens zwei bestimmte Kriterien hin optimiert.

Dieser Cluster an möglichen Sequenzen wurde in zwei Bereiche geteilt, einen nativen und einen nicht-nativen Bereich. Alle nicht-nativen Sequenzen haben folgende Aminosäuren gegenüber dem Wildtyp (1vii) ausgetauscht: A4D, E24W, AYW32-34EKG und weisen laut Vorhersage eine etwas andere Tertiärstruktur auf als der Wildtyp.

Der Cluster aus vier nativen und fünf nicht-nativen Sequenzen sollte aus Primärsequenzen bestehen, die alle erstens strukturell stabiler sein sollten als der Wildtyp und zweitens eine Liganden-zugängliche Schleife besitzen (vgl. Abb. 1.6). Diese beiden Kriterien wurden gewählt, da besagte Eigenschaften oft in praktischen Anwendungen gewünscht werden.

Gronwald *et al.* haben 2008 die bereits 2005 von Hohm *et al.* erwähnte G34L-Mutante von HP36 über NMR-Spektroskopie als tatsächlich stabiler gegenüber dem Wildtyp bestätigt, indem er und Kollegen die über NMR-Spektroskopie ermittelte Struktur mit der für die G34L-Mutante vorhergesagten Struktur verglichen haben. Diese G34L-Mutante stammte aus dem Teil des oben genannten Clusters der nativen HP36-Mutanten.

Auffällig war jedoch der Teil des Clusters der nicht-nativen Mutanten, welche bis zu neun ausgetauschte Aminosäuren gegenüber dem Wildtyp HP36 besaßen. Eine dieser nicht-nativen Mutanten, die Achtfachmutante r1g11i1, wurde daraus ausgewählt. Diese sollte sehr gut in wässriger Umgebung löslich sein und bot dadurch eine geeignete Voraussetzung, um *in vitro* untersucht werden zu können. Das Peptid r1g11i1 sollte als Test dienen, ob EMOS auch für den Teil des nicht-nativen Clusters korrekte Vorhersagen getroffen hat.

Aus der über EMOS erhaltenen Primärsequenz für rlg11i1 wurde eine stabile 3D-Struktur vorhergesagt. Diese 3D-Struktur (new2) sollte nun auf ihre Korrektheit hin überprüft werden, um beurteilen zu können, ob EMOS´ Vorhersage fehlerfrei war. Für dieses Vorhaben wurde mut8 (*in vitro*-Äquivalent zur vorhergesagten Primärsequenz rlg11i1) chemisch synthetisiert (siehe Kap. 2.2) und über NMR-Spektroskopie charakterisiert, um das Resultat der Strukturaufklärung mit dem der bioinformatischen Vorhersage (new2) zu vergleichen.

## 3.1. NMR-Spektroskopie von mut8

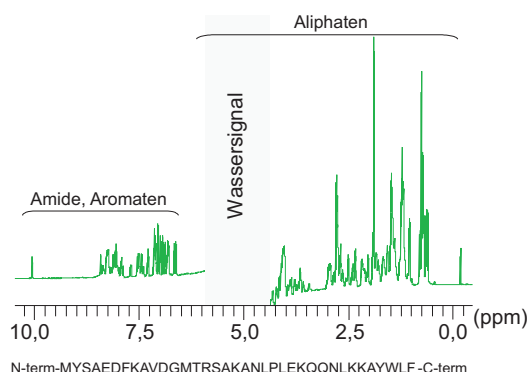
Die nachfolgenden Kapitel beschreiben die Auswertung der aufgenommenen 1D- und homonuklearen 2D-NMR-Spektren ( $^1\text{H}$ -, COSY-, TOCSY- und NOESY-Spektren). Mit Hilfe dieser Spektren war es möglich, zum einen die Sekundärstruktur-Elemente von mut8 zu bestimmen und zum anderen eine Tertiärstruktur für mut8 zu berechnen. Dafür müssen alle im Spektrum auftauchenden Resonanzen zugeordnet werden (siehe Kap. 3.1.1) und sowohl die Konnektivitäten über NOEs (siehe Kap. 3.1.2) als auch der chemische Verschiebungsindex (CSI) (siehe Kap. 3.1.2) und die Diederwinkel  $\phi$  und  $\psi$  (siehe Kap. 3.1.2) bestimmt werden. Die dabei gewonnenen Erkenntnisse fließen in die anschließende Berechnung der Tertiärstruktur von mut8 mit ein (siehe Kap. 3.1.3).

### 3.1.1. Spektrenaufnahme und -zuordnung

#### 1D-Spektroskopie des chemisch synthetisierten Peptids

Die NMR-spektroskopische Charakterisierung des synthetischen Peptids mut8 wurde mit der Aufnahme eines 1D- $^1\text{H}$ -Spektrums begonnen (siehe Kap. 2.5.3). Durch die Aufnahme eines solchen Spektrums lassen sich zwei wichtige Informationen gewinnen. Erstens kann die Reinheit der Synthese von mut8 überprüft werden und zweitens lässt sich eine Aussage darüber treffen, ob mut8 strukturiert vorliegt oder nicht. Für die Aufnahme des 1D- $^1\text{H}$ -Spektrums wurde mut8 in einer Konzentration von 1 mM in 50 mM KPi-Puffer, pH 6,5 ( $\text{H}_2\text{O}:\text{D}_2\text{O}/90:10$ ) gelöst, die Lösung in ein NMR-Röhrchen überführt und bei 25 °C im NMR-Spektrometer (700 MHz, Bruker Biospin) gemessen (vgl. Kap. 2.5.3). Abbildung 3.1 zeigt das prozessierte 1D-Spektrum und die passende Primärsequenz von mut8.

Zu erkennen sind Signale im Bereich der aliphatischen Protonen zwischen 0,5 und 5 ppm und die der aromatischen und Amid-Protonen zwischen 6,5 und 10 ppm. Das Signal bei -0,19 ppm repräsentiert das Methylsignal des zugegebenen Frequenzstandards 2,2-Dimethyl-2-silapentanon-5-sulfonsäure (DSS). Das Signal bei 10 ppm resultiert aus der Anwesenheit des  $\epsilon$ -Indol Protons aus dem Tryptophanring von W34. Die Tatsache, dass nur ein W34-Signal zu erkennen ist, weist darauf hin, dass nur eine Population mut8 existiert. Somit kann an der Stelle die Reinheit der Synthese bestätigt werden. Das Wassersignal wurde im Spektrum der Übersicht halber in der Abbildung ausgespart. Die gute Verteilung der Signale weist auf eine ausgeprägte Wechselwirkung benachbarter Atome miteinander hin, die zu einer hohen Disperion über das gemessene Frequenzband führen. Diese Beobachtung legt den Schluss nahe, dass es sich bei dem



**Abbildung 3.1.:** 1D- $^1\text{H}$ -Spektrum des synthetisch hergestellten mut8. Die Signale von aromatischen und Amidprotonen sowie von Aliphatenprotonen sind durch Klammern markiert. In der Mitte des Spektrums befindet sich bei 4,79 ppm ein Signal, das durch das in der NMR-Probe befindliche Wasser erzeugt wird. Ebenfalls abgebildet ist die Primärsequenz von mut8.

synthetischen mut8 um ein strukturiertes Peptid handelt. Fehlende Amidprotonen-Signale zwischen 8,5 und 9,5 ppm, welche typisch für  $\beta$ -Faltblattstränge sind, lassen vermuten, dass mut8 eher  $\alpha$ -helikale Anteile hat [93, 97].

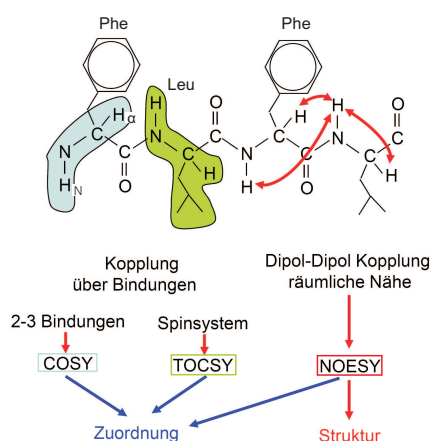
### Homonukleare 2D-NMR des chemisch synthetisierten Peptids

Die Aufnahme von COSY-, TOCSY- und NOESY-Spektren liefert für ein kleines Peptid wie mut8 in der Regel genügend Informationen (siehe Abb. 3.2), um sowohl die Sekundärstrukturelemente zu bestimmen als auch die Tertiärstruktur zu berechnen. Daher wurden im Anschluss an die Aufnahme des 1D- $^1\text{H}$ -Spektrums besagte zweidimensionale, homonukleare Spektren aufgenommen (siehe Kap. 2.5), sodass die Signale aus den Spektren den einzelnen Protonen des Peptids zugeordnet werden können (siehe Kap. 2.5.5).

In Abbildung 3.3 sind die drei übereinandergelegten COSY-, TOCSY- und NOESY-Spektren zu sehen. Mit einem grauen Kasten ist der *Fingerprint*-Bereich des TOCSY-Spektrums eingerahmt. Dieser zeigt Signale möglicher Kopplungen der  $\text{H}_\text{N}$  mit den jeweiligen  $\text{H}_{\alpha,\beta,\gamma,\delta,\epsilon}$  – falls vorhanden – der eigenen Aminosäure.

Nach Aufnahme aller 2D-homonuklearen Spektren bietet es sich an, mit der Analyse des COSY-Spektrums zu beginnen, da dort die übersichtlichste Anzahl an Signalen zu finden ist. In einem COSY reicht der Magnetisierungstransfer lediglich über maximal drei Bindungen zwischen zwei koppelnden Protonen (siehe Kap. 2.5.5). Daher sollten im Idealfall im *Fingerprint*-Bereich eines COSY-Spektrums alle  $\text{H}_\alpha$ - $\text{H}_\text{N}$ -Kopplungen (Entfernung von 3 Bindungen zwischen  $\text{H}_\alpha$  und  $\text{H}_\text{N}$ ) zu erkennen sein. So kann durch Auszählen aller im COSY-*Fingerprint*-Bereich sichtbaren COSY-Signale (blau) ermittelt werden, ob alle 36 Aminosäuren des Peptids mut8 im Spektrum sichtbar sind.

In der COSY-*Fingerprint*-Region müssten nun bei einer Anzahl von 36 Aminosäuren des mut8 auch mindestens 34 COSY-Signale zu erkennen sein. Die Anzahl 34 (statt



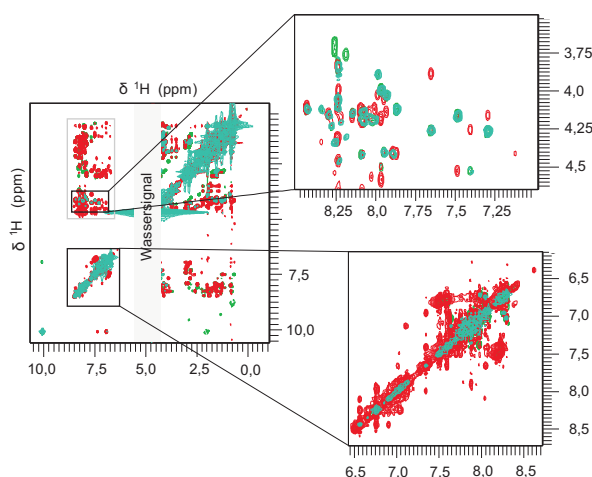
**Abbildung 3.2.:** Schematische Darstellung eines Tripeptids aus zwei Phenylalaninen und einem Leucin und die jeweiligen Kopplungen, die sich im COSY-, TOCSY- und NOESY-Spektrum ergeben würden. Im COSY-Spektrum werden Signale sichtbar, welche durch die Kopplung zwischen zwei Protonen über maximal 3 Bindungen zustande kommen (blau). Das TOCSY-Spektrum zeigt die Signale, die durch die Kopplung zwischen Protonen eines Spinsystems (einer Aminosäure) zustande kommen (grün). Alle COSY-Signale sind somit auch TOCSY-Signale. Die im NOESY-Spektrum sichtbaren Signale (rot) lassen sich darauf zurückführen, dass Protonen, die maximal 5 Å voneinander entfernt sind, miteinander koppeln. Verändert nach [98].

36) ergibt sich daraus, dass es für P22 aufgrund des fehlenden H<sub>N</sub> kein entsprechendes H<sub>N</sub>-H<sub>α</sub>-Kopplungssignal gibt und dass die H<sub>N</sub>-aliphatischen Korrelationen für die erste Aminosäure in den meisten Fällen nicht zu sehen ist.

Es ließen sich 26 COSY-Signale in der COSY-*Fingerprint*-Region des Spektrums identifizieren (siehe Abb. 3.3, oben rechts, Anzahl der blauen Signale), dabei entspricht die COSY-*Fingerprint*-Region einem Teil der TOCSY-*Fingerprint*-Region (grauer Kasten). Die Tatsache, dass acht COSY-Signale nicht sichtbar sind, lässt sich darauf zurückführen, dass mehrere COSY-Signale überlagern können und daher keine klare Zuordnung stattfinden kann. Dies ist z. B. im Bereich um 8,0/4,2 ppm der Fall (siehe Abb. 3.3, oben rechts).

In einem an die Analyse des COSYs folgenden Schritt wurde der Bereich der H<sub>N</sub>-H<sub>N</sub>-Kopplungen zwischen 6,5 und 8,5 ppm im NOESY analysiert (siehe Abb. 3.3, unten rechts). In einem NOESY-Spektrum ergeben alle Kopplungen zwischen zwei Protonen ein Signal (einen NOE (*Nuclear Overhauser Effect*)), die – ob über Bindungen oder über den Raum – maximal 5 Å voneinander entfernt sind. Somit liefert ein NOESY-Spektrum Informationen über die Tertiärstruktur eines Peptids bzw. Proteins.

In dem in Abb. 3.3 sichtbaren, unten rechts befindlichen Ausschnitt der überlagerten COSY-, TOCSY-, und NOESY-Spektren lassen sich deutlich mindestens 26 NOEs erkennen, welche für eine Strukturierung des Proteins sprechen. Zusätzlich wurde eine mut8-Probe lyophilisiert und in 100 % D<sub>2</sub>O resuspendiert. Davon wurde anschließend ein NOESY aufgenommen. Für die Protonen, welche mit dem <sup>2</sup>H austauschen,

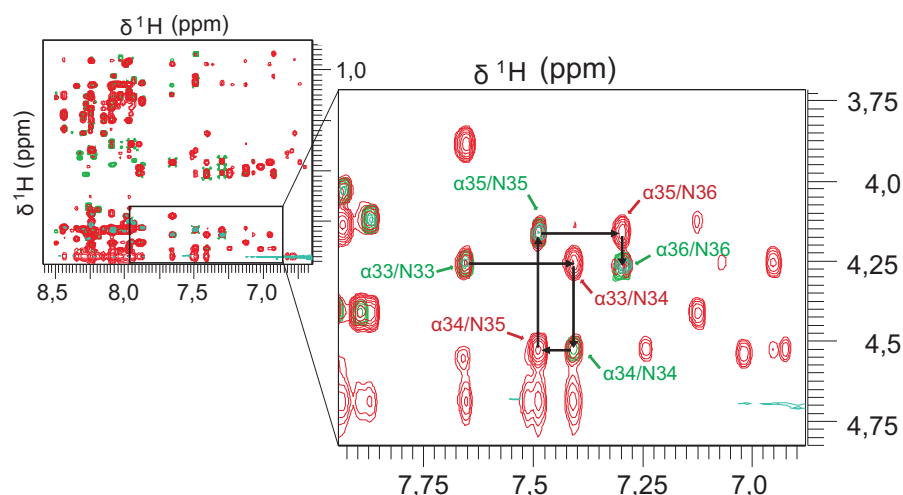


**Abbildung 3.3.:** Die drei übereinandergelegten 2D-Spektren und Ausschnitte aus der *Fingerprint*-Region des COSYs (oben rechts) sowie dem Bereich der  $H_N$ - $H_N$ -Kopplungen (unten rechts). Die Signale des COSY-Spektrums sind in blau, die des TOCSY-Spektrums in grün und die des NOESY-Spektrums in rot dargestellt. Der graue Kasten markiert den TOCSY-*Fingerprint*-Bereich, bei etwa 4,67 ppm befindet sich das Wassersignal. Der besseren Übersicht wegen wurde oben rechts die Intensität der NOESY-Signale herabgesetzt, unten rechts die der COSY-Signale.

sind keine Signale mehr im Spektrum sichtbar. Sind alle oder so gut wie alle Signale im NOESY verschwunden, so deutet dies auf eine Dynamik des Peptids hin. Tatsächlich konnten im TOCSY-*Fingerprint*-Bereich und in der  $H_N$ - $H_N$ -Region des Spektrums nur vier deutliche Signale und vier schwächere zugeordnet werden. Diese ließen sich zumeist auf intraresiduale Kopplungen zurückführen: Intraresiduale Kopplungen zwischen  $7H_\beta/7H_\delta$ ,  $33H_\alpha/33H_\delta$ ,  $34H_\beta/34H_\delta$  und die interresiduale Kopplung zwischen  $34H_\delta/35H_\delta$  waren als schwächere Signale im Spektrum zu sehen, intraresiduale Kopplungen zwischen  $2H_\beta/2H_\delta$ ,  $2H_\delta/2H_\epsilon$ ,  $33H_\beta/33H_\epsilon$  und  $33H_\delta/33H_\epsilon$  waren als starke Signale im Spektrum erkennbar.

### Zuordnung der Signale

Nach Überlagerung von COSY-, TOCSY- und NOESY-Spektren konnte mit der Zuordnung der Signale und dem *Chaintracing* begonnen werden. Das *Chaintracing* stellt einen Teil des Zuordnungsverfahrens dar und erlaubt es, die  $H_\alpha$ - $H_N$ -Signale von Protonen aus COSY/TOCSY-Spektren den jeweiligen Aminosäuren einer Peptidkette eindeutig zuzuordnen. Die genaue Zuordnungsstrategie über das *Chaintracing* lässt sich Kap. 2.5.5 entnehmen. Abbildung 3.4 zeigt exemplarisch die Zuordnungsstrategie, die für mut8 verwendet wurde, ausgehend von dem  $H_\alpha$ - $H_N$ -Signal der Aminosäure Y33 bis zum letzten Signal  $H_\alpha$ - $H_N$ -Kopplung der Aminosäure F36. Zu sehen sind hierbei der TOCSY-*Fingerprint*-Bereich (links), welcher in Abbildung 3.3 mit einem grauen Kasten umrandet ist, sowie ein Ausschnitt der entsprechenden Region (rechts). In rot



**Abbildung 3.4.:** Überlagerung von TOCSY- und NOESY-Spektrum des synthetischen mut8-Peptids. Zu sehen ist der *Fingerprint*-Bereich der in Abb. 3.3 übereinandergelegten Spektren (links oben) und ein Ausschnitt der  $H_\alpha$ - $H_N$ -Region (rechts) mit *Chaintracing* – von Aminosäure Y33 bis F36. TOCSY-Signale sind in grün dargestellt, NOESY-Signale in rot und COSY-Signale in blau.

sind wiederum die Kreuzresonanzen des NOESY- und in grün die Kreuzresonanzen des TOCSY-Spektrums dargestellt.

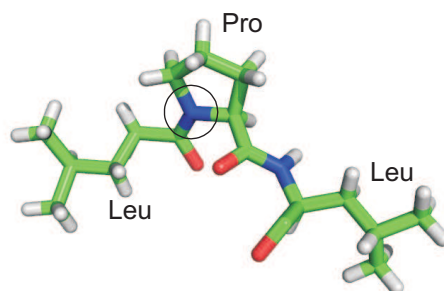
### Das Chaintracing

Jedes im Ausschnitt von Abbildung 3.4 sichtbare und beschriftete TOCSY-Signal (grün) ergibt sich aus der Kopplung zwischen dem  $H_\alpha$  ( $\alpha$ ) und dem  $H_N$  (N) der eigenen Aminosäure  $i$ . Für jedes dieser Signale findet man ein entsprechendes Signal im COSY-Spektrum, welches zur eindeutigen Identifizierung der  $H_\alpha$ - $H_N$ -Kopplung dient.

Neben den  $H_{\alpha i}$ - $H_{Ni}$ -Signalen finden sich in den COSY- und TOCSY-Spektren weitere Kopplungen. Über letztere lassen sich Spinsysteme identifizieren, die zwar einzelnen Aminosäurentypen zugeordnet werden können, jedoch liefern sie keine Informationen über die Sequenzposition dieser Reste. Mit Hilfe eines zusätzlichen NOESY-Spektrums lassen sich diese Informationen jedoch gewinnen. Da im NOESY-Spektrum zwei Protonen, die maximal 5 Å voneinander entfernt liegen (über den Raum oder über Bindungen), einen NOE hervorrufen, sieht man zusätzlich zu den entsprechenden  $H_\alpha$ - $H_N$ -Signalen des TOCSYs bzw. COSYs auch ein NOESY-Signal zwischen dem  $H_\alpha$  ( $\alpha$ ) der eigenen Aminosäure  $i$  und dem  $H_N$  (N) der Aminosäure  $i+1$ , also der C-terminal benachbarten Aminosäure.

Geht man von dem COSY-Signal für  $\alpha 33/N33$  aus, so muss der entsprechende NOE ( $H_{\alpha 33}/H_{N34}$ ), der die Verbindung zu dem Signal  $\alpha 34/N34$  herstellt, die Frequenz des  $C_\alpha$ -Protons der Aminosäure 33 und die Frequenz des Amidprotons von Aminosäure 34 tragen. Dieser NOE liegt daher im Spektrum im Schnittpunkt einer waagrechten Linie,





**Abbildung 3.5.:** Darstellung von Prolin innerhalb einer Peptidkette aus Leu-Pro-Leu (erstellt mit Pymol). Eingekreist ist das Stickstoffatom, welches kein Proton mehr besitzt. Dadurch fehlen die Signale der Kopplungen zwischen dem  $H_N$  und den Ringprotonen.

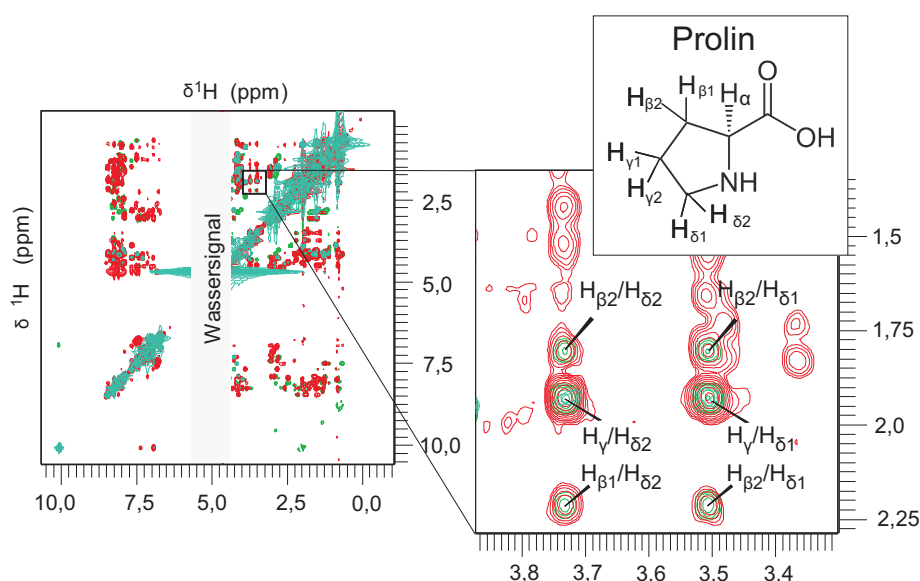
die durch das COSY-Signal  $\alpha 33/N33$  läuft und einer vertikalen Linie, die durch das entsprechende Signal von  $\alpha 34/N34$  geht. Diese Art der Verknüpfung der Spinsysteme mit Hilfe von COSY/TOCSY- und NOESY-Spektren und einer bekannten Primärsequenz wird *Chaintracing* genannt. Analog zu dem hier angeführten Beispiel wurden alle übrigen Aminosäuren von S3 bis L21 und L23 bis F36 zugeordnet. Insgesamt konnten so etwa 92 % aller  $H_N$ -Rückgratprotonen zugeordnet werden.

Unterbrochen wurde diese klassische *Chaintracing*-Prozedur dabei durch die Aminosäure Prolin P22, welche aufgrund des fehlenden Amidprotons (siehe Abb. 3.5) keinen  $H_{\alpha 21}/H_{N22}$ -NOE im NOESY-Spektrum ausbilden kann.

### Die Zuordnung des Prolins

Obwohl Prolin nicht über das *Chaintracing*-Verfahren zugeordnet werden kann, konnten sowohl das  $H_{\alpha}$  als auch die  $H_{\beta,\gamma,\delta}$  ermittelt werden. Charakteristisch für Prolin ist, dass die Kopplungen der Ringprotonen ein charakteristisches Muster aus drei parallel zueinander angeordneten Signalen im Aliphatenbereich des Spektrums ergeben (siehe Abb. 3.6, rechts). Somit ließen sich die  $H_{\delta}$ -Frequenzen bei 3,50 ppm ( $H_{\delta 1}$ ) und 3,73 ppm ( $H_{\delta 2}$ ) eindeutig erkennen und zuordnen.

Bei dem TOCSY-Signal um 1,93/3,73 ppm befindet sich noch ein COSY-Signal. So konnte geschlussfolgert werden, dass es sich dabei um die  $H_{\gamma 2}/H_{\delta 2}$ -Kopplung handeln muss, da zwischen allen anderen Protonen zum  $H_{\delta}$  mehr als drei Bindungen liegen. Dasselbe gilt für die parallel dazu liegende  $H_{\gamma}/H_{\delta 1}$ -Kopplung. Die übrigen TOCSY-Signale aus diesem Ausschnitt konnten somit den  $H_{\beta}/H_{\delta}$ -Kopplungen zugeordnet werden. Die Zuordnung des übrig gebliebenen  $H_{\alpha}$  ergibt sich nun daraus, dass sich für die Kopplung zwischen diesem Proton und dem  $H_{\beta}$  ein COSY-Signal um 4,32/1,80 ppm ergeben müsste, was auch tatsächlich der Fall war. Um das Prolin sequentiell einbinden zu können, wurden die Signale bei 7,96/3,50 und 7,96/3,73 ppm verwendet. Diese entsprechen der Kopplung zwischen dem  $H_N$  des L21 und den  $H_{\delta}$  des P22.



**Abbildung 3.6.:** Überlagerte 2D-Spektren (vgl. 3.3) und ein Ausschnitt aus der Aliphaten-Region (rechts). Zu sehen sind im Spektrum-Ausschnitt die charakteristischen Signale der Kopplungen der Ringprotonen des Prolins P22 (Protonen nach IUPAC bezeichnet).

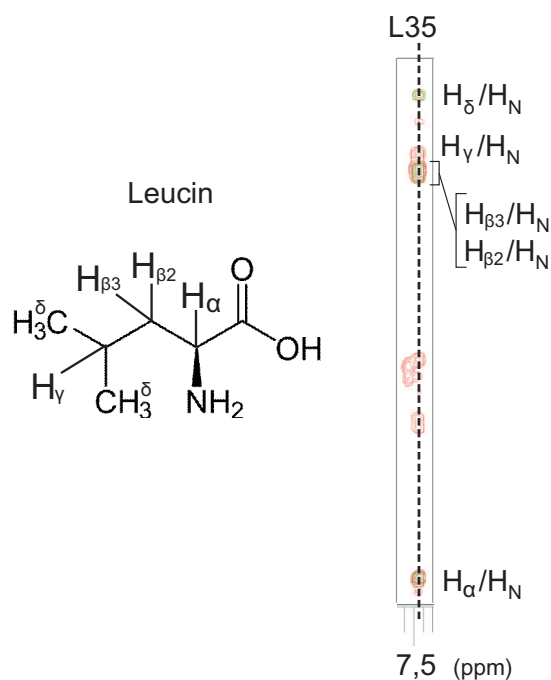
### Die Zuordnung von aliphatischen Seitenketten

Ebenfalls zugeordnet wurden soweit möglich alle Seitenketten-Protonen, darunter die  $H_\beta$ ,  $H_\gamma$ ,  $H_\delta$  und  $H_\epsilon$ , deren Signale im Bereich zwischen 0,8 und 3,8 ppm zu finden sind [90]. Diese koppeln im TOCSY-Spektrum meist mit dem  $H_N$  der jeweiligen Aminosäure, woraus sich charakteristische Muster im Spektrum ergeben (siehe Kap. 2.5.5).

Abbildung 3.7 zeigt exemplarisch die Zuordnung von Seitenkettenprotonen anhand von Aminosäure L35 (vgl. auch Abbildung 3.4) über das charakteristische Muster für Leucin im *Fingerprint*-Bereich des TOCSY-Spektrums (siehe Abb. 2.4). Dort lassen sich zusätzlich zum  $H_{\alpha i}/H_{N i}$ -Signal (hier entspricht i Aminosäure L35) die Kopplungen zwischen den beiden  $H_{\beta i}$  und dem  $H_{N i}$  und zwischen den  $H_{\gamma i}/H_{\delta i}$  und  $H_{N i}$  zuordnen.

Da die  $H_\delta$  der beiden Methylgruppen chemisch äquivalent sind, lässt sich im TOCSY-Spektrum nur ein Signal erkennen. Die weiterhin in Abbildung 3.7 sichtbaren reinen NOESY-Signale (rot) stellen Kopplungen zwischen Seitenkettenprotonen anderer Aminosäuren und dem  $H_N$  von L35 dar.

Über die hier beschriebene Zuordnungsstrategie konnten 100 % aller  $H_\alpha$  und 74 % aller  $H_\beta$  sowie etwa 59 % aller übrigen aliphatischen, nicht-Aminogruppen-haltigen Seitenketten zugeordnet werden.



**Abbildung 3.7.:** Vollständige Zuordnung von L35, Ausschnitt aus der TOCSY-Fingerprint-Region (rechts) und Leucin als Strukturmodell mit Bezeichnung aller Protonen nach der IUPAC-Nomenklatur (links).

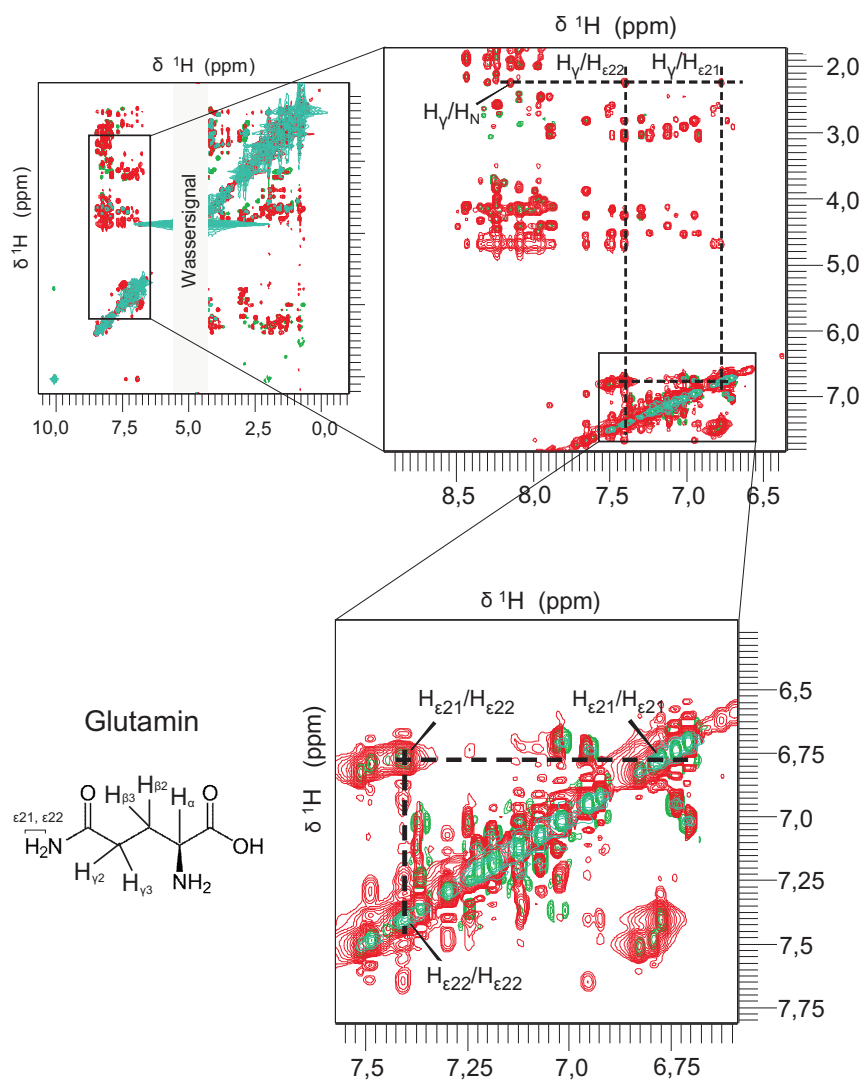
### Die Zuordnung von Amino-Protonen

Die Signale für Aminoprotonen von Lysin, Asparagin, Glutamin und Arginin und deren Kopplungen zu aliphatischen Nachbarn sind in den 2D-Spektren in der *Fingerprint*-Region zwischen 6,5 und 8,5 ppm lokalisiert. Abbildung 3.8 zeigt exemplarisch die Zuordnung der Aminoprotonen von Q26.

Für ein Glutamin erwartet man im TOCSY-Spektrum Signale für die Kopplungen zwischen dem  $H_{\epsilon 21}$  und dem  $H_{\epsilon 22}$  sowie NOEs von beiden zu den benachbarten  $H_{\gamma}$ . Dabei kann es sein, dass für die  $H_{\gamma}$  nur ein Signal zu sehen ist, da beide Protonen chemisch äquivalent sind. Dies ist hierbei auch der Fall.

Auf der Frequenz der  $H_{\gamma}$  bei 2,27 ppm findet sich ein NOE bei 2,27/6,77 (siehe Abb. 3.8 (rechts oben)), der die Kopplung zwischen den  $H_{\gamma}$  und  $H_{\epsilon 21}$  repräsentieren. Einen entsprechenden NOE von den  $H_{\gamma}$  zu dem  $H_{\epsilon 22}$  findet sich bei 2,27/7,40 ppm. Verifizieren lässt sich die Zuordnung dadurch, dass es für das Signal bei 6,77/7,40 ppm sowohl im NOESY- als auch im TOCSY- und COSY-Spektrum ein entsprechendes Signal gibt, der die Kopplung der beiden  $H_{\epsilon}$  untereinander repräsentiert.

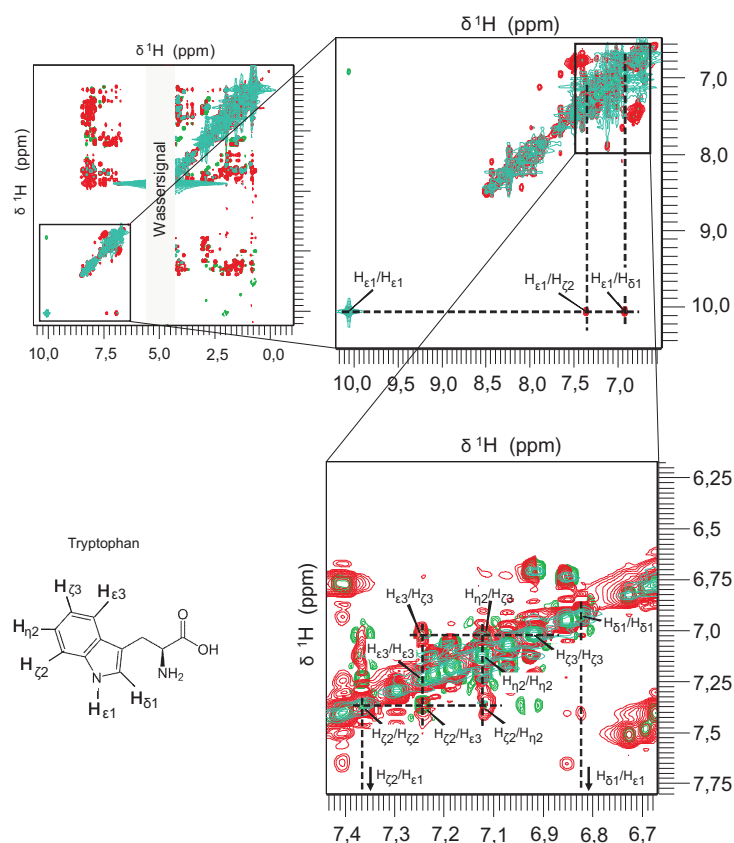
Über diese Strategie konnten anschließend auch die Seitenketten- $H_N$ s von N20, Q27 und N28 zugeordnet werden; insgesamt wurden damit 40 % dieser Protonen eindeutig bestimmt.



**Abbildung 3.8.:** Die drei übereinandergelagerten 2D-Spektren (vgl. Abb. 3.3) und ein Ausschnitt (oben), welcher den ersten Zuordnungsschritt der Aminoprotonen von Q26 zeigt. Unten ist der oben rechts markierte Bereich vergrößert dargestellt, links davon befindet sich das Glutamin mit Bezeichnung der Protonen nach IUPAC-Nomenklatur.

### Die Zuordnung von Seitenketten-Ringprotonen

Eine besondere Stellung für die Zuordnungsstrategie nehmen die Ringprotonen der Seitenketten von Phenylalanin, Tyrosin, Tryptophan und Histidin ein, da sie jeweils charakteristische Signalmuster ergeben. Diese sind in den aufgenommenen 2D-Spektren zwischen 6,5 und 8,7 ppm sichtbar. Eine markante Ausnahme bildet das Indolproton  $H_{\epsilon 1}$  des Tryptophans, da dessen Signal immer bei etwa 10 ppm zu sehen und somit sehr leicht zu bestimmen ist.



**Abbildung 3.9.:** Die drei übereinandergelegten 2D-Spektren (vgl. Abb. 3.3) und ein Ausschnitt (oben), welcher den ersten Zuordnungsschritt der Ringprotonen von W34 zeigt. Der untere Vergrößerungsausschnitt zeigt die Zuordnung der übrigen Ringprotonen, welche untereinander koppeln. Links davon ist das Tryptophan mit Bezeichnung der Protonen nach IUPAC-Nomenklatur dargestellt.

Im Folgenden ist exemplarisch die Zuordnung der Ringprotonen des W34 beschrieben (siehe Abb. 3.9): Ausgehend von dem bei 10,06 ppm erkennbaren Signal des  $H_{\epsilon 1}$  lässt sich die Kopplung bei 10,06/6,90 ppm als TOCSY/COSY-Signal von  $H_{\epsilon 1}/H_{\delta 1}$  zuordnen, da das  $H_{\delta 1}$  das einzige Proton ist, das weniger als vier Bindungen von  $H_{\epsilon 1}$

entfernt ist und so ein COSY-Signal hervorrufen kann.

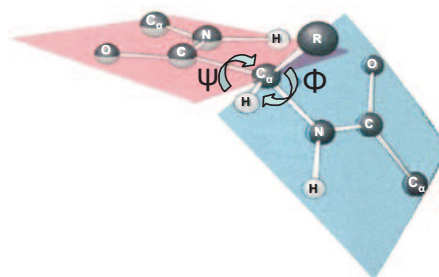
Ein weiteres NOESY-Signal auf der Frequenzebene des  $H_{\epsilon 1}$  bei 10,06/7,36 ppm lässt sich als  $H_{\epsilon 1}/H_{\zeta 2}$  bestimmen, da diese beiden Protonen in einem planaren Tryptophanring räumlich näher beieinander sind als alle übrigen Protonen im Ringsystem (siehe Abb. 3.9, rechts oben). Von dem zugeordneten  $H_{\zeta 2}$  lassen sich nun alle weiteren Ringprotonen zuordnen. So ergibt sich für die Kopplung zwischen dem  $H_{\zeta 2}$  und dem  $H_{\eta 2}$  ein COSY-Signal, dasselbe gilt auch für die Kopplung zwischen  $H_{\eta 2}/H_{\zeta 3}$  und  $H_{\zeta 3}/H_{\epsilon 3}$ , so dass diese eindeutig zugeordnet werden können (siehe Abb. 3.9, unten rechts). Analog zu dieser Strategie wurden auch die Ringprotonen des Y2, F7, Y33 und F36 und damit insgesamt 63 % der Ringprotonen zugeordnet.

Zusammenfassend sind in Tabelle 3.1 alle zugeordneten Protonen der 36 Aminosäuren von mut8 aufgelistet.

### 3.1.2. Sekundärstrukturanalyse

Nachdem die Signale aller Aminosäuren von mut8 mit Hilfe der drei homonuklearen COSY-, TOCSY- und NOESY-Spektren zugeordnet werden konnten, wurde eine Sekundärstrukturanalyse von rlg11 durchgeführt.

Anhand der Verteilung und Intensität bestimmter NOEs sowie aus den chemischen Verschiebungen von Resonanzen lassen sich direkt Aussagen über den Gehalt und das Vorkommen von Sekundärstrukturelementen treffen. Die chemische Umgebung eines Atoms, welcher Teil des Proteinrückgrats bzw. unmittelbar daran gebunden ist, wird durch die Sekundärstruktur an dieser Stelle beeinflusst, da durch diese die Bindungsgeometrie des Rückgrats bestimmt wird. Aus diesem Grund lässt sich eine Korrelation der chemischen Verschiebungen dieser Atome mit der Sekundärstruktur feststellen [99].



**Abbildung 3.10.:** Darstellung der Diederwinkel  $\phi$  und  $\psi$  innerhalb einer Peptidkette. Durch die partielle Doppelbindung ist die Peptidbindung selbst starr, sodass eine Drehung des Peptidrückgrats nur über die Diederwinkel erfolgen kann. Verändert nach [100].

Weiterhin ist es möglich, über die Diederwinkel eines Proteins oder Peptids Aussagen über dessen Sekundärstruktur zu treffen. Da die C-N-Bindung des Peptidrückgrates nur eingeschränkt drehbar ist, sind lediglich Rotationen um die N- $C_{\alpha}$ - und  $C_{\alpha}$ -C-Bindungen möglich. Diese Drehungen werden durch besagte Diederwinkel  $\phi$

**Tabelle 3.1.:** Auflistung aller zugeordneten Protonen mit Hilfe der 2D-Spektren. Die Bezeichnung der Atome erfolgt nach der IUPAC-Nomenklatur. Mit eingefügt sind vorausgreifend die Zuordnungen der Stickstoff-Resonanzen (siehe Kap. 3.4.1).

AS	H <sub>N</sub>	H <sub>α</sub>	H <sub>β</sub>	H <sub>γ</sub>	H <sub>δ</sub>	H <sub>ε</sub>	N
1		3,89	1,96	2,40	–		
2		4,53	2,92	–	7,02	6,71	
3	8,26	4,34	3,73/3,74	–	–	–	119,17
4	8,34	4,11	1,31	–	–	–	126,42
5	8,24	4,06	1,87/1,77	2,12	–	–	118,48
6	7,95	4,41	2,47	–	–	–	120,51
7	7,89	4,41	3,06/2,94	–	7,12	7,22	120,20
8	7,87	4,12	1,63/1,61	1,26		2,86	122,45
9	8,01	4,18	1,29	–	–	–	124,83
10	7,97	3,99	2,00	0,83	–	–	118,79
11	8,23	4,44	2,58	–	–	–	123,21
12	8,23	3,85	–	–	–	–	108,81
13	8,09	4,43	1,98	2,51/2,42	–		119,62
14	8,07	4,20	4,12	1,26	–	–	115,34
15	8,27	4,23	1,78/1,69	1,53	3,06	–	123,34
16	8,19	4,30	3,77	–	–	–	116,51
17	8,07	4,21	1,11	–	–	–	122,45
18	8,21	4,18	1,60	1,30		2,27	126,14
19	8,04	4,18	1,27	–	–	–	124,10
20	8,22	4,59	2,72/2,60	–	6,79/7,49	–	117,55
21	7,96	4,49	1,525	1,43	0,83	–	123,38
22	–	4,32	2,21/1,80	1,93	3,50/3,73	–	–
23	8,24	4,08	1,54/1,47	1,47	0,79	–	122,48
24	8,44	4,11	1,88/1,93	2,17	–	–	120,65
25	8,08	4,14	1,76/1,64	1,34	1,57	2,90	119,89
26	8,15	4,14	1,93/2,02	2,27	–	6,77/7,40	120,51
27	8,29	4,19	1,98/1,87	2,25	–	6,78/7,4	120,55
28	8,32	4,56	2,73/2,64	–	6,83/7,51	–	119,10
29	8,03	4,17	1,54	1,48	0,74/0,80	–	122,10
30	8,08	4,12	1,75/1,65	1,34		2,87	121,27
31	7,93	4,02	1,58/1,5	1,20		2,79	121,76
32	7,99	3,89	0,89	–	–	–	124,62
33	7,65	4,26	2,86/2,81	–	6,95	6,74	118,37
34	7,41	4,51	3,07/3,03	–	6,92	10,06	120,76
35	7,49	4,18	1,28/1,22	1,11	0,68	–	123,52
36	7,29	4,29	3,03/2,80	–	7,07	7,19	124,93
	H <sub>ε3</sub>	H <sub>ζ2</sub>	H <sub>ζ3</sub>	H <sub>η2</sub>	H <sub>ζ</sub>		
7	–	–	–	–	7,18		
34	7,24	7,36	7,02	7,12	–		

und  $\psi$  beschrieben (siehe Abb. 3.10). Aus sterischen Gründen sind innerhalb einer Peptidkette nur bestimmte Kombinationen aus Diederwinkeln möglich, wodurch sich für Sekundärstrukturelemente wie  $\alpha$ -Helices und  $\beta$ -Faltblätter charakteristische Winkelkombinationen ergeben. Diese Winkelkombinationen werden zur Berechnung der Tertiärstruktur eines Peptids oder Proteins eingesetzt.

Im Folgenden werden die Ergebnisse verschiedener Analysewege (Analyse nach Wüthrich, Analyse nach Wishart und Vorhersage der Diederwinkel) zur datenbasierten Sekundärstrukturbestimmung von mut8 gezeigt.

### Analyse nach Wüthrich

Das Vorhandensein bestimmter NOEs innerhalb des NOESY-Spektrums erlaubt es, ein strukturelles Element in einem Peptid oder Protein als  $\alpha$ -Helix oder  $\beta$ -Faltblatt zu deklarieren [92]. Sequentiell aufeinanderfolgende NOEs zwischen Amidprotonen der Aminosäure  $i$  und Aminosäure  $i+2$  ( $d_{NN}(i, i+2)$ ) sowie von  $d_{\alpha N}(i, i+3)$ ,  $d_{\alpha N}(i, i+4)$  und  $d_{\alpha\beta}(i, i+3)$  sind charakteristisch für  $\alpha$ -helikale Bereiche.

So waren 24 sequentielle NOEs zwischen einem  $H_\alpha$  der Aminosäure  $i$  und dem C-terminal folgenden Amidproton der Aminosäure  $i+1$  ( $d_{\alpha N}(i, i+1)$ ) eindeutig zu finden. Zwei dieser NOEs ( $d_{\alpha N}(4, 5)$  und  $d_{\alpha N}(6, 7)$ ) waren nur ungenau bestimmbar, da sie von anderen Signalen überlagert wurden.

Danach folgten 14 sequentielle NOEs zwischen einem Amidproton  $i$  und dem C-terminal folgenden Amidproton  $i+1$  ( $d_{NN}(i, i+1)$ ). Davon waren wiederum 3 aufgrund von Überlagerungen mit anderen Signalen nur ungenau bestimmbar. 7 NOEs zwischen dem Amidproton der Aminosäure  $i$  und dem Amidproton der C-terminal gelegenen Aminosäure  $i+2$  ( $d_{NN}(i, i+2)$ ) wurden ermittelt, davon waren abermals 3 nur ungenau bestimmbar, da Signale überlagerten.

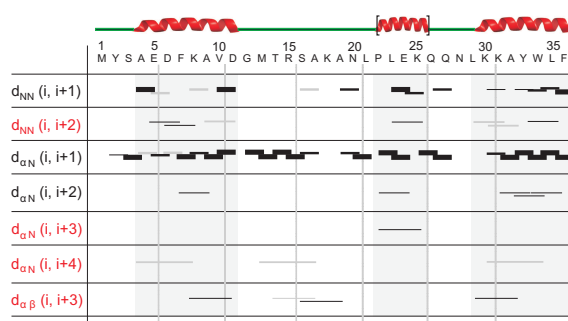
Für die Kopplung zwischen dem  $H_\alpha$  der Aminosäure  $i$  und dem Amidproton der Aminosäure  $i+2$  wurden 5 Konnektivitäten gefunden, welche alle eindeutig zuzuordnen waren ( $d_{\alpha N}(i, i+2)$ ). Weiterhin wurde nur ein NOE zugeordnet, der der Kopplung zwischen dem  $H_\alpha$  der Aminosäure 22 und dem Amidproton der Aminosäure 25 entspricht ( $d_{\alpha N}(i, i+3)$ ). Dieser konnte eindeutig bestimmt werden.

Es wurden 3 NOEs zwischen dem  $H_\alpha$  der Aminosäure  $i$  und dem Amidproton der Aminosäure  $i+4$  ermittelt ( $d_{\alpha N}(i, i+4)$ ), welche allerdings ungenau zu bestimmen waren, da auch dort Signale überlagerten.

Abschließend wurden für die Kopplung zwischen dem  $H_\alpha$  der Aminosäure  $i$  und dem  $H_\beta$  der Aminosäure  $i+3$  vier NOEs gefunden ( $d_{\alpha\beta}(i, i+3)$ ), davon konnte aufgrund von Signalüberlagerungen ein NOE nur ungenau zugeordnet werden.

Eine bildliche Darstellung der Konnektivitäten von mut8 ist in Abbildung 3.11 zu sehen. Dort sind potentielle  $\alpha$ -helikale Bereiche grau unterlegt sowie als *cartoon*-Modell über der Peptidsequenz von mut8 abgebildet. Aufgrund der NOE-Daten bilden die Aminosäuren 5 bis 11 und 29 bis 36  $\alpha$ -Helices, während die  $\alpha$ -Helix im Bereich von Aminosäuren 22 bis 25 weniger gut definiert ist.





**Abbildung 3.11.:** Darstellung der Konnektivitäten innerhalb von mut8. Die ganz linke Spalte der Tabelle zeigt die Konnektivitäten zwischen den Protonen der einzelnen Aminosäuren in Balkendarstellung. Je stärker ein Balken ist, desto intensiver ist der entsprechende NOE. Graue Balken stehen für nicht eindeutig definierbare NOEs. In rot geschrieben (linke Spalte) sind die NOEs, welche ausschlaggebend für die Existenz von  $\alpha$ -Helices sind. Über der Tabelle ist die mut8-Sequenz im Einbuchstabencode zu sehen, darüber die *cartoon*-Darstellung des Peptids mit den  $\alpha$ -Helix-Anteilen; in eckige Klammern gefasst ist die weniger gut definierte  $\alpha$ -Helix. Die Bereiche der potentiellen Helices sind grau unterlegt.

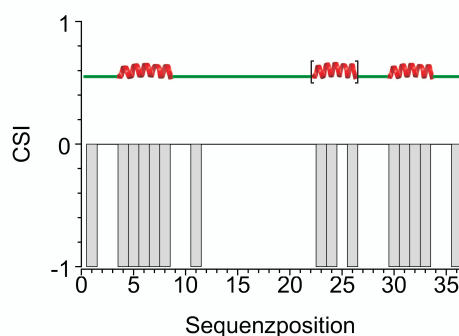
### Analyse nach Wishart – Der chemische Verschiebungsindex (CSI)

Neben der Analyse der Sekundärstrukturelemente auf Basis von NOEs, lassen sich auch aufgrund der chemischen Verschiebungen Aussagen über mögliche Sekundärstrukturen machen.

Die Analyse nach Wishart erlaubt es, anhand der unterschiedlichen chemischen Verschiebungen der  $\alpha$ -Protonen eines Peptids oder Proteins relativ zu einem Standard wie 2,2-Dimethyl-2-silapentan-5-sulfonsäure (DSS)  $\alpha$ -Helices oder  $\beta$ -Faltblätter vorherzusagen. Dazu werden die Resonanzfrequenzen von  $\alpha$ -Protonen einer Aminosäure x in Proteinen/Peptiden mit den entsprechenden Frequenzen in identischer Aminosäure x eines ungefalteten und sekundärstrukturlosen Hexapeptids GGXAGG, das als Referenzwert dient, verglichen. Während die Resonanzfrequenzen der  $\alpha$ -Protonen von  $\alpha$ -Helices gegenüber den Frequenzen in den Hexapeptiden hochfeldverschoben sind, sind die Resonanzfrequenzen der  $\alpha$ -Protonen von  $\beta$ -Faltblättern tieffeldverschoben [93].

Weicht der im Protein/Peptid gemessene Wert um mehr als  $\pm 0,1$  ppm vom Referenzwert ab, so erhält das entsprechende  $H_\alpha$  den Wert  $\pm 1$ . Ist die Abweichung kleiner als  $0,1$  ppm, wird dem entsprechenden  $H_\alpha$  der Wert 0 zugeordnet. Dieses Verfahren wird als Ermittlung des Chemischen Verschiebungsindex (CSI) bezeichnet. Folgen 3 oder mehr  $+1$ -Werte, die nicht durch eine  $-1$  unterbrochen werden, hintereinander, dann handelt es sich bei dem Bereich um ein  $\beta$ -Faltblatt, folgen 4 oder mehr  $-1$ -Werte, die nicht durch eine  $+1$  unterbrochen werden, hintereinander, so liegt in dem Bereich eine  $\alpha$ -Helix vor.

Abbildung 3.12 zeigt die Darstellung des CSI für das mut8-Peptid. Hier wurden anhand der vorangegangenen Zuordnung aller Protonen (siehe Tab. 3.1) die  $H_\alpha$ -Werte



**Abbildung 3.12.:** Darstellung des chemischen Verschiebungsindex (CSI) für mut8. Zu sehen ist die Verteilung der ermittelten CSI-Werte für das Peptid. Die  $\alpha$ -helikalen Bereiche liegen zwischen den Aminosäuren 4 und 8 bzw. 30 und 33, ein weniger gut definierter  $\alpha$ -helikaler Bereich liegt zwischen Aminosäuren 23 und 26. Über der Verteilung ist eine *cartoon*-Darstellung des Peptids mit den  $\alpha$ -Helix-Anteilen dargestellt; in eckige Klammern ist die weniger gut definierte  $\alpha$ -Helix gefasst.

von mut8 mit denen aus der Referenztable (Tabelle II aus Wishart *et al.*, 1992 [101]) verglichen, nachdem sie um das Referenzsignal (DSS) korrigiert wurden. In Abbildung 3.12 ist zu erkennen, dass von Aminosäuren 4 bis 8 und von Aminosäuren 30 bis 33 eine  $\alpha$ -Helix vorhergesagt wird, ein dritter  $\alpha$ -Helix-Bereich liegt zwischen den Aminosäuren 23 und 26. Letzter Kap. ist weniger gut definiert als die beiden anderen Bereiche, da die Viererkette der -1-Werte von einem 0-Wert unterbrochen wird.

### Vorhersage der Diederwinkel mit TALOS

Mit Hilfe des Software-Pakets TALOS [102] lassen sich aufgrund der chemischen Verschiebung der von Resonanzfrequenzen bestimmten Kernen Vorhersagen für die Torsionswinkel  $\phi$  und  $\psi$  des Peptids erstellen.

Chemische Verschiebungen von Rückgratprotonen in Peptiden und Proteinen sind sehr sensitiv bezüglich lokaler Konformationen; homologe Proteine zeigen dabei ähnliche chemische Verschiebungen. Diese Information wird genutzt, um innerhalb einer Datenbank nach Triplets von aufeinanderfolgenden Aminosäuren mit chemischen Verschiebungen und Sequenzähnlichkeit zu suchen, die die beste Übereinstimmung mit dem Zieltriplett innerhalb der eigenen Proteinsequenz – hier mut8 – liefern. Die Datenbank enthält sowohl chemische Verschiebungen für C- als auch für N- und  $H_\alpha$ -Atome für 20 Proteine, für welche eine hochaufgelöste Kristallstruktur vorhanden ist. Werden Übereinstimmungen gefunden, so ordnet TALOS die Torsionswinkel der gespeicherten Aminosäuren eines Proteins aus dem Datenbankeintrag den entsprechenden Aminosäuren des Triplets aus mut8 zu. TALOS lieferte nach Einlesen der chemischen Verschiebungen aller Protonen aus dem Fingerprint- Bereich von mut8 insgesamt 68 Winkelinformationen, davon 34 für die  $\phi$ - und 34 für die  $\psi$ -Winkel.

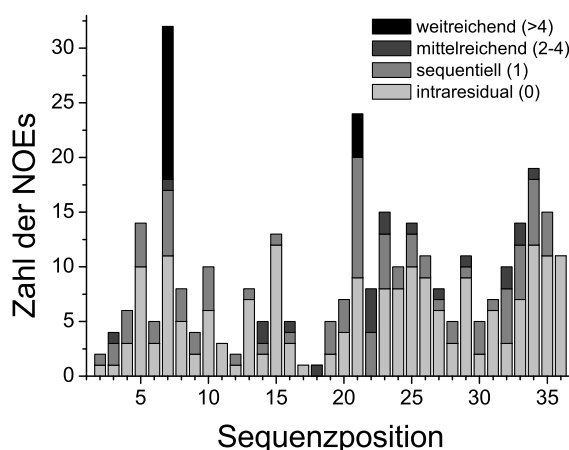
### 3.1.3. Tertiärstrukturberechnung

Die stabile Tertiärstruktur ist ausschlaggebend für die biologische Funktion eines Peptids oder Proteins, daher ist es wichtig, bei allen *in silico* generierten Peptiden auf das Vorhandensein einer stabilen dreidimensionalen Struktur zu achten. Das Peptid mut8 sollte laut Vorhersage ebenfalls eine stabile Tertiärstruktur besitzen. Der Nachweis der vorhergesagten stabilen Faltung des Peptids über NMR-Spektroskopie würde das bioinformatisch erhaltene Resultat verifizieren und die Vorhersage somit als korrekt klassifizieren.

Nach der Sekundärstrukturanalyse von mut8 wurde die Tertiärstruktur des Peptids berechnet als mut8 bezeichnet. Zur Berechnung wurden sowohl die aufgrund der über NOE-Intensitäten ermittelten Abstände (siehe Kap. 2.5.5) zwischen den Protonen als auch die über TALOS vorhergesagten Winkelinformationen eingesetzt. Verwendet wurden alle Winkelvorschläge bis auf die für M13 und T14 ( $\psi$ -Winkel) sowie für T14 und R15 ( $\phi$ -Winkel), da durch diese während der finalen Tertiärstrukturberechnung zu hohe energetische Verletzungen auftraten.

Die Berechnung der Tertiärstruktur von mut8 wurde mit Hilfe eines *simulated annealing*-Algorithmus aus dem Programmpaket Xplor-NIH durchgeführt. Hierbei wird aus der Primärstruktur des Peptids *in silico* eine gestreckte Peptidkette generiert. Den Atomen der Kette werden dann stochastisch Geschwindigkeiten zugeordnet, deren Maxwellverteilung einer hohen Temperatur entspricht (z. B. 1000 K). Anschließend erfolgt eine Moleküldynamiksimulation unter Lösungen der Newton'schen Bewegungsgleichungen mit nachfolgender Energieminimierung. Die Zielfunktion für die Minimierung enthält Terme für die experimentell erhobenen Einschränkungen (NOEs, Diederwinkel), idealisierte kovalente Geometrien und Nicht-Bindungs-Kontakte. Durch stufenweises Abkühlen auf Raumtemperatur nimmt das Molekül *in silico* langsam einen gefalteten Zustand ein, der am Simulationsende der Tertiärstruktur des Makromoleküls entspricht. Durch Variation der Zuordnung der Geschwindigkeiten auf die Atome der primär erzeugten Molekülkette erhält man unterschiedliche Ausgangszustände, die zu leicht unterschiedlichen Lösungsstrukturen führen. Daher ergibt sich bei der Berechnung einer NMR-Struktur eine Schar von Strukturen, deren Abweichungen voneinander von der Anzahl und Qualität der experimentell ermittelten Einschränkungen abhängt. Zur Tertiärstrukturberechnung werden Verfahren zur Energieminimierung genutzt, die Moleküldynamik im kartesischen Raum verwenden [76, 77]

Alle gewonnenen NOE-Distanzbeschränkungen wurden in intraresiduale (i,i), sequentielle (i,i+1), mittelweite ( $\leq(i,i+4)$ ) und weitreichende ( $>(i,i+4)$ ) unterteilt und bis auf die intraresidualen Beschränkungen in Xplor-NIH zur Strukturberechnung eingesetzt. Letztere wurden erst am Ende der Kalkulationen hinzugefügt, um eine Verbesserung der Struktur durch Verringerung der Energie und somit optimalere Faltung zu erreichen. Tabelle 3.2 zeigt die Anzahl und Art der verwendeten NOE-Beschränkungen. Von den 194 intraresidualen NOEs (i,i) waren 14 Rückgrat-zu-Rückgrat (RzR)-NOEs, 124 Rückgrat-zu-Seitenketten (RzS)-NOEs und 56 Seitenketten-zu-Seitenketten (SzS)-NOEs. Die 95 verwendeten sequentiellen NOEs (i,i+1) lassen sich in 40 RzR-, 46 RzS- und 9 SzS-NOEs einteilen. Für Aminosäuren 11, 17 und 18 wurden keine eindeutigen sequentiellen NOEs gefunden. Von den 25 eingesetzten mittelweiten NOEs ( $\leq(i,i+4)$ ),

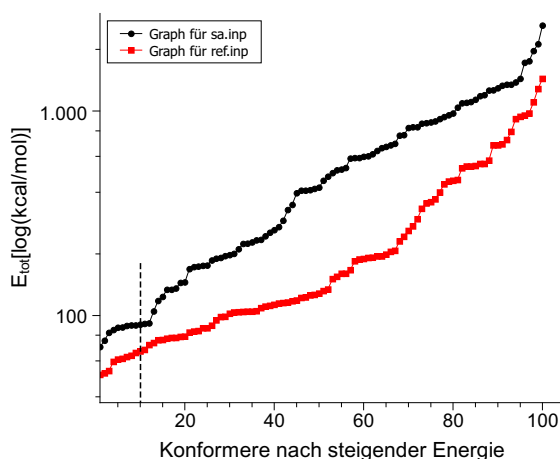


**Abbildung 3.13.:** Anzahl und Art der zur Tertiärstrukturberechnung eingesetzten NOE-Beschränkungen pro Aminosäure. Die hellgrauen Balken markieren die intraresidualen Kopplungen, die mittelgrauen Balken die sequentiellen Kopplungen, die dunkelgrauen die mittelreichenden und die schwarzen die weitreichenden Kopplungen.

i+4)) waren 12 RzR-, 7 RzS- und 5 SzS-NOEs. Für Aminosäuren 2, 4, 8 bis 13, 17, 19 bis 21, 24, 26, 28, 30, 31 und 35 wurden keine eindeutigen mittelweit reichenden NOEs gefunden. 19 eindeutige NOE-Beschränkungen wurden für die weitreichenden ( $>(i,i+4)$ ) gefunden. Darunter waren 11 RzS und 8 SzS-NOEs. Die sequentiellen NOEs waren zumeist zwischen Aminosäuren 7 $\rightarrow$ 8 (6x), 21 $\rightarrow$ 22 (11x) und 34 $\rightarrow$ 35 (7x) zu finden, bei den mittelweiten NOEs traten zumeist Kopplungen zwischen Aminosäuren 22 $\rightarrow$ 24/25 (4x), 23 $\rightarrow$ 25/26/27 (3x) und 33 $\rightarrow$ 35 (3x) auf. Für die weit reichenden NOEs wurden Kopplungen zwischen Aminosäuren 7 $\rightarrow$ 30/33/34/35/36 (15x) und 21 $\rightarrow$ 26/29 (4x) aufgeführt. Abbildung 3.13 zeigt zusammenfassend die Anzahl und Art der NOE-Beschränkungen, welche zur Tertiärstrukturberechnung eingesetzt wurden.

Zuerst wurde eine *mut8.psf*-Datei über das Protokoll *generate.inp* erzeugt, dann wurde über das Protokoll *generate\_template.inp* eine Strukturdatei generiert (siehe Kap. 2.6.3, Anhang A). Beide Dateien sowie die NOE-Beschränkungen und die Torsionswinkel (TALOS) wurden zur Berechnung der Tertiärstruktur mit Xplor-NIH über das *sa\_mut8.inp*-Protokoll aus der XPLOR-Programmbibliothek eingesetzt (siehe Kap. 2.6.3). Als experimentelle Parameter wurden insgesamt 333 NOE-Beschränkungen zur Strukturberechnung verwendet, davon waren 194 intraresiduale, 95 sequentielle, 25 mittelreichende und 19 weitreichende Beschränkungen. Hinzu kamen 64 dihedrale Beschränkungen. Das *sa\_mut8.inp*-Protokoll enthält einen *Simulated-Annealing* (SA)-Algorithmus.

Es wird mit einer Starttemperatur von 1000 K begonnen und eine Molekulardynamiksimulation bestehend aus 6000 Schritten durchführt, um dann in einer Abkühlphase auf 100 K von 3000 Schritten langsam eine Endtemperatur von 100 K zu erreichen. Die Schritte des SA-Protokolls wurden 100x durchgeführt; so wurde eine Strukturschar von insgesamt 100 möglichen Konformeren erzeugt. Zur Verfeinerung wurde das *ref.inp*-Protokoll von Nilges *et al.*, 1991 verwendet [103], welches eine langsame-



**Abbildung 3.14.:** Energieprofil der erhaltenen Strukturenscharen nach Verwendung des *sa.inp* (schwarz)- und des *ref.inp* (rot)-Protokolls. Dargestellt sind die Gesamtenergien der einzelnen Konformere, die ersten 10 mit der geringsten Energie wurden als Ensemble ausgewählt.

re Abkühlphase bestehend aus 4000 Schritten beinhaltet. Somit können u. a. aufgrund des durch das vorangegangene kürzere Abkühlungsprotokoll eventuell entstandene thermodynamisch ungünstige Konformere korrigiert werden (siehe Kap. 2.6.4). In Abbildung 3.14 sind die Energieverteilungen der Strukturenscharen für das *sa.inp*- und das *ref.inp*-Protokoll dargestellt. Deutlich zu erkennen ist, dass nach Einsatz des Verfeinerungs-Protokolls die Energien signifikant sinken.

Im Anschluss an die Verfeinerung wurde aus den 100 energetisch minimierten Konformeren ein finales Ensemble von 10 Strukturen (*mut8*) ausgewählt, die weder Verletzungen der NOE-Distanzbeschränkungen noch Verletzungen der über TALOS erhaltenen dihedralen Beschränkungen aufwiesen und die kleinsten Energien innerhalb der 100er-Strukturenschar besaßen. Je geringer die Energie eines Systems, in dem Fall eines Peptids, desto größer ist in den meisten Fällen die Wahrscheinlichkeit, dass dieses eine korrekte, der Realität nahekommende Konformation aufweist.

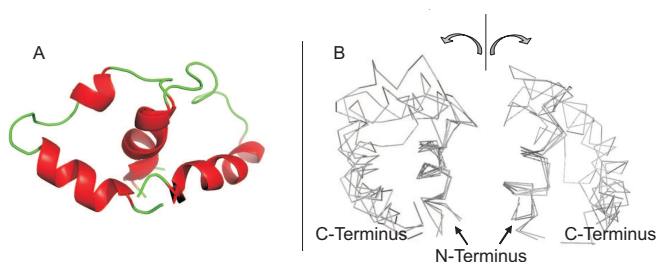
Detaillierte Angaben zu den Charakteristika des Ensembles befinden sich in Tabelle 3.2.

Das Ensemble wurde in PyMol (siehe Kap. 2.6.5) eingelesen und entlang der  $C_{\alpha}$ -Atome der Aminosäuren 3 bis 11 (erste  $\alpha$ -Helix) übereinandergelegt (siehe Abb. 3.15). Interessanterweise lässt sich spätestens an der Stelle feststellen, dass die eindeutige Bestimmung einer Tertiärstruktur von *mut8* nicht möglich ist, da Xplor-NIH augenscheinlich nicht in der Lage war, anhand der eingesetzten Beschränkungen eindeutige Konformationen für sowohl den N- als auch C-Terminus des Peptids zu berechnen.

Daher wurde auf die Erstellung einer gemittelten Struktur des Ensembles verzichtet und stattdessen dieses in zwei verschiedene Cluster geteilt (siehe Abb. 3.15, B). Die N-terminale  $\alpha$ -Helix und die C-terminale Helix sind in sich sehr gut definiert, die mittlere Helix dagegen nicht (siehe Abb. 3.15, A). Diese ist in 60 % der Ensemble-

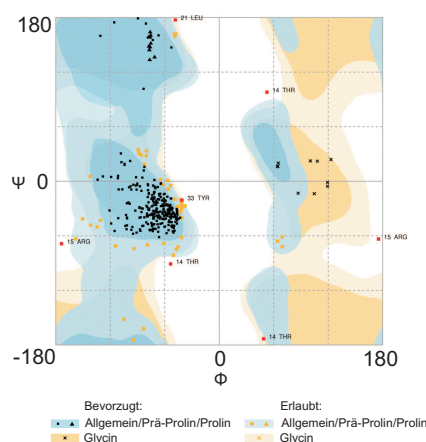
**Tabelle 3.2.:** Strukturelle Statistik für das energieminierte Ensemble aus 10 Strukturen berechnet für mut8 mit Xplor-NIH.

<b>NOE-Abstandsbeschränkungen</b>	
Intraresidual (i,i)	194
Sequentiell (i,i+1)	95
Mittelweit ( $\leq(i,i+4)$ )	25
Weitreichend ( $>(i,i+4)$ )	19
<b>Sonstige Beschränkungen</b>	
Dihedrale Beschränkungen	64
<b>Ramachandranplot-Regionen (% aller AS)</b>	
favorisiert	85
erlaubt	13
außerhalb liegend	2
<b>Root mean square deviation (RMSD)(Å)</b>	
Aminosäuren 3 bis 11	0,98
1. Cluster, AS 1 bis 36	2,75
2. Cluster, AS 1 bis 36	1,96
<b>Durchschnittsenergien (kcal/mol)</b>	
Bindungsstreckung	1,99
Winkelbiegung	42,16
<i>Improper</i>	6,44
Van der Waals	4,49
NOE	3,63
Dieder	1,24
Total	59,95
<b>Durchschnittlicher Sekundärstrukturgehalt (%)</b>	
$\alpha$ -Helix	54
Schleife/ungeordnet	46



**Abbildung 3.15.:** Darstellung der für mut8 berechneten Strukturen des Ensembles mut8. (A) *cartoon*-Darstellung zweier möglicher mut8-Strukturen des Ensembles. Die erste  $\alpha$ -Helix sowie die letzte sind gut definiert, die mittlere dagegen nicht. (B) Aufteilung des Ensembles in zwei Cluster. 1: 5 mögliche mut8-Strukturen des Ensembles mit einem nach „links“ gerichteten Mittelteil und C-Terminus. 2: 4 mögliche mut8-Strukturen des Ensembles mit einem nach „rechts“ gerichteten Mittelteil und C-Terminus. Eine der 10 Strukturen ist weder nach „links“ noch nach „rechts“ orientiert (nicht gezeigt). Alle Strukturen wurden entlang der C $_{\alpha}$ -Atome der ersten  $\alpha$ -Helix (AS 3 bis 11) überlagert.

Strukturen zu finden, in 40 % befindet sich dort eine Schleife. Relativ zur überlagerten ersten  $\alpha$ -Helix gesehen, ist der übrige Teil des Peptids in seiner Position nicht genauer definiert. Die RMSD lag bei 0,98 Å für die erste Helix, bei 2,75 Å für den ersten Cluster (Aminosäuren 1 bis 36) und bei 1,96 Å für den zweiten Cluster (Aminosäuren 1 bis 36). Mittels RamPage [1] (siehe Kap. 2.6.5) wurde ein Ramachandranplot erstellt (siehe Abb. 3.16). Etwa 85 % aller Aminosäuren liegen in favorisierten Bereichen, 13 % in erlaubten Bereichen und 2 % in Ausreißerbereichen.



**Abbildung 3.16.:** Ramachandranplot des Ensembles für mut8 (RamPage [1]).

Trotz der schlecht definierbaren Tertiärstruktur von mut8 konnten alle Sekundärstrukturbereiche gut bestimmt werden, somit ergab sich nach Einlesen des Ensembles

in MoleMan [74] (siehe Kap. 2.6.5) ein  $\alpha$ -Helix-Anteil von 54 % und ein Anteil von 46 % an Schleifen bzw. ungeordneten Bereichen.

Während sich zwar die Sekundärstrukturelemente von mut8 gut bestimmen ließen, war die Berechnung einer eindeutigen Tertiärstruktur nicht möglich. Dies kann verschiedene Ursachen haben: Zum einen könnten NOEs bei der Zuordnung übersehen worden sein, was jedoch ausgeschlossen werden kann. Zum anderen waren weitreichende NOEs teilweise sehr schwach, nicht eindeutig zuzuordnen und widersprüchlich.

Dies könnte ein Hinweis darauf sein, dass mut8 entgegen der anfänglichen Annahme (Vorliegen von nur einem Konformer, siehe Kap. 3.1.1) in verschiedenen stabilen Zuständen vorliegt. Ist dies der Fall, so können diese schnell oder langsam ineinander übergehen, das Peptid also eine bestimmte interne Dynamik besitzen.

Um dieser Vermutung nachzugehen, müssen Dynamikmessungen über NMR erfolgen. Diese Messungen erfordern  $^{15}\text{N}$ -markiertes Peptid, da nun keine homonuklearen Spektren mehr aufgenommen werden, sondern heteronukleare  $^1\text{H}$ - $^{15}\text{N}$ -Spektren.

## 3.2. Klonierung von mut8

Während die strukturellen Informationen aufgrund des kleinen Molekulargewichts von mut8 aus den homonuklearen Spektren gewonnen werden konnten, benötigte man für die Aufnahme von HSQCNOE (hetNOE)-Spektren zur Bestimmung der Dynamik ein  $^{15}\text{N}$ -isotopenangereichertes Peptid. Letzteres konnte aus Kostengründen nicht synthetisch hergestellt werden und wurde daher rekombinat in *E. coli* exprimiert. Dazu wurde in einem ersten Schritt eine entsprechende DNA-Sequenz entworfen, durch PCR vervielfältigt und in einen Vektor integriert (vgl. Kap. 2.3.1, siehe Abb. 2.3).

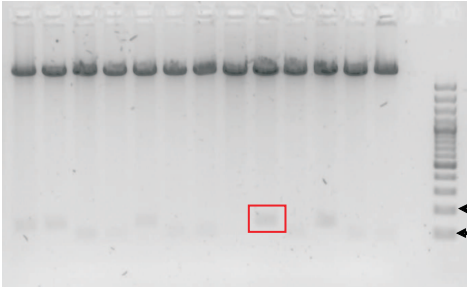
Dafür wurden 0,7 pmol eines Annealing-Ansatzes (Endkonzentration 300 pmol/  $\mu\text{l}$  einer 1:1-Mischung der synthetischen mut8-DNA-Sequenzen (mut8- und revers komplementäre Sequenz) (Microsynth)) über PCR mit Hilfe der Primer 646\_F und 647\_R (siehe Kap. 2.3.1) amplifiziert.

Sowohl das Amplifikat als auch der Zielvektor pET41b\_mod (siehe Kap. 2.3.1) wurden mit ApaI und XhoI restringiert und im Anschluss daran in einem Verhältnis von 5:1 (Amplifikat:Vektor) miteinander ligiert. Das Peptid liegt nach diesem Klonierungsschritt als GST (Glutathion-S-Transferase)-Konstrukt vor (C-terminaler tag), um das Peptid nach der Expression über eine GSH (Glutathion)-Affinitätssäule reinigen zu können. Eine Schnittstelle für die PreScission-Protease zwischen tag und Peptid sollte im Anschluss an die Reinigung die proteolytische Spaltung und Freisetzung von mut ermöglichen. Die genaue Beschreibung der Klonierungsstrategie findet sich im Kap. 2.3.1.

Zum Nachweis einer erfolgreichen Klonierung wurden *E. coli* XL10 Gold-Zellen mit 0,15  $\mu\text{g}$  des Ligationsansatzes transformiert, dieser auf einer Agarplatte ausplattiert und über Nacht bei 37 °C inkubiert. Es waren am Folgetag insgesamt 13 Klone zu sehen, welche jeweils in 8 ml LB-Medium angezogen wurden, sodass die Vektor-DNA dieser Klone extrahiert (siehe Kap. 2.3.2) und 0,1  $\mu\text{g}$  davon mit den Restriktionsenzymen ApaI und XhoI geschnitten wurden.

Im 1,5 %igen Agarosegel waren ein 100 bp- und ein etwa 6000 bp-Fragment zu sehen





(siehe Abb. 3.17), was den korrekten, zu erwartenden Größen entspricht. Das Kleinere der beiden Fragmente stellt das klonierte GST\_mut8-Fragment dar, das Größere den Vektor pET41b\_mod. 80 ng/  $\mu$ l der übrigen DNA-Minipräparation wurden zur Sequenzierung geschickt (GATC) und das positive Klonierungsergebnis bestätigt (siehe Abb. 3.18). Der Expressionsvektor wird im Folgenden als pET41b\_mod\_GST\_mut8 bezeichnet.

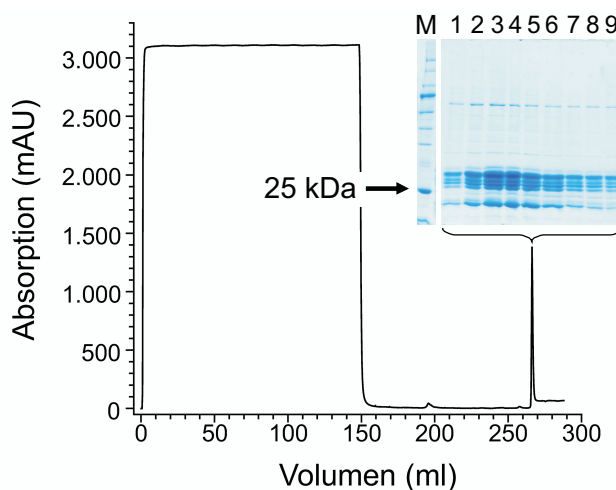
Kontrolle mut8	GGTGGTGGCGACCATCTCCAAAAATCGGATGGTTCAACTAGTGGTTCTGGTCATCACCAT	780
	<div style="display: flex; align-items: center; justify-content: space-between;"> <div style="text-align: center;"> <div style="width: 20px; height: 10px; background-color: black; margin: 0 auto;"></div> <div style="text-align: left;">← GST-tag</div> </div> <div style="text-align: center;"> <div style="width: 100px; height: 10px; background-color: red; margin: 0 auto;"></div> <div style="text-align: left;">PreScission-Sequenz</div> </div> <div style="text-align: center;"> <div style="width: 20px; height: 10px; background-color: black; margin: 0 auto;"></div> <div style="text-align: left;">Apal</div> </div> </div>	
Kontrolle mut8	CACCATCACTCCGCGGGTCTGGAAGTTCTGTTCCAGGGGCCCATGTATAGCGCAGAAGAT	24 840
Kontrolle mut8	TTTAAAGCAGTTGATGGTATGACCCGTAGCGCAAAAGCAAAATCTGCCGCTGGAAAAACAG	84
	TTTAAAGCAGTTGATGGTATGACCCGTAGCGCAAAAGCAAAATCTGCCGCTGGAAAAACAG	900
Kontrolle mut8	CAGAATCTGAAAAAAGCATATTGGCTGTTTAACTCTCGAG	124
	<div style="display: flex; align-items: center; justify-content: space-between;"> <div style="text-align: center;"> <div style="width: 20px; height: 10px; background-color: black; margin: 0 auto;"></div> <div style="text-align: left;">Stop</div> </div> <div style="text-align: center;"> <div style="width: 20px; height: 10px; background-color: black; margin: 0 auto;"></div> <div style="text-align: left;">Xho1</div> </div> </div>	960
Kontrolle mut8	CCACTAATTGATTAATACCTAGGCTGCTAAACAAAGCCCGAAAGAAGCTAGTCCC	1015

### 3.3. Expression und Reinigung von mut8

Nach der erfolgreichen Klonierung von mut8 wurde das Konstrukt in *E. coli* transformiert und das Peptid exprimiert. Anschließend erfolgte die Reinigung über GSH-Affinitäts- und Gelfiltrationschromatographie. Da die Aufnahme von  $^{15}\text{N}$ -NMR-Spektren zur Bestimmung von Dynamik des Peptids (siehe Kap. 3.4) geplant war, musste das rekombinante Peptid mit dem Isotop  $^{15}\text{N}$  angereichert werden. Daher wurden zur Anzucht der *E. coli*-Zellen ein Minimal-Wachstumsmedium (siehe Kap. 2.3.3) mit zugesetztem  $^{15}\text{N}$ -markiertem Ammoniumchlorid als Stickstoffquelle verwendet.

#### 3.3.1. Expression und GSH-Affinitätschromatographie

Der Vektor pET41b\_mod\_GST\_mut8 wurde in den Expressionsstamm *E. coli* Rosetta (DE3) transformiert und in 4 l M9-Minimalmedium mit zugesetztem  $^{15}\text{N}$ -markiertem Ammoniumchlorid (1 g/l) auf eine OD600 von 0,8 angezogen. Anschließend erfolgte die Induktion der Expression von GST\_mut8 durch Zugabe von 0,2 mM IPTG; die Hauptkultur wurde bei 28 °C weiterhin schüttelnd über Nacht inkubiert. Am nächsten Tag wurde die Hauptkultur abzentrifugiert und das Pellet in Waschpuffer mit 1 mM PMSF und 0,2 mM EDTA als Proteaseinhibitoren resuspendiert. Im Anschluss daran wurden die Zellen durch Zugabe von 5 mg Lysozym und Verwenden des Mikrofluidizers aufgeschlossen. Die so aufgeschlossenen Zellen wurden in der Ultrazentrifuge abzentrifugiert und der filtrierte Überstand auf eine 1 ml GSH-Säule (GE-Healthcare) gegeben (vgl. Kap. 2.4.1, Kap. 2.4.2). Um den Verlauf der Reinigung von GST\_mut8 verfolgen zu können, wurden die Fraktionen nach der Affinitätschromatographie auf ein SDS-Gel aufgetragen.



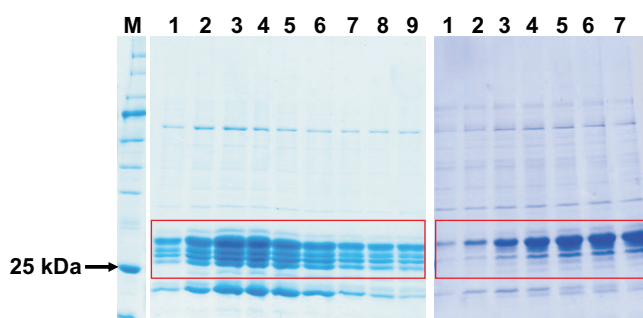
**Abbildung 3.19.:** Elutionsverlauf der GSH-Reinigung mit 12,5 %igen SDS-Gel nach Coomassie-Färbung. Bei Volumen 0 bis 150 ml nach Beladen der Säule werden die nicht bindenden Proteine im Durchfluss registriert. Nach 265 ml beginnt die Elution des GST\_mut8-Fusionsproteins und endet bei etwa 270 ml.

Auf dem SDS-Gel (Abb. 3.19) sind links (Spur M) 4  $\mu$ l des Markers aufgetragen (NEB), danach folgen jeweils 5  $\mu$ l der einzelnen über die GSH-Säule gereinigten Fraktionen in den Spuren 1 bis 9. In letzteren Spuren ist eine deutliche Hauptbande bei etwa 32 kDa zu erkennen, was der zu erwartenden Größe aus GST-*tag* (28 kDa) und mut8 (4,2 kDa) entspricht. Neben dieser Hauptbande sind noch weitere Banden zwischen 25 und 30 kDa zu erkennen, bei denen es sich um C-terminale Abbauprodukte des GST\_mut8-Fusionsproteins handelt.

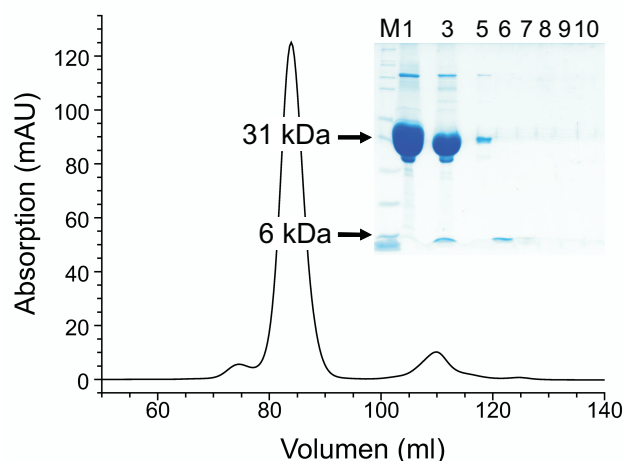
### 3.3.2. Optimierung der Expression und Reinigung

Der bereits erwähnte Abbau des Fusionsproteins GST\_mut8 (vgl. Kap. 3.3.1) musste gestoppt oder vermindert werden, um zu vermeiden, dass bei den nachfolgenden NMR-spektroskopischen Untersuchungen lediglich ein Fragment des eigentlichen mut8-Peptides vorliegt, was die strukturelle und dynamische Charakterisierung des Peptids verfälschen könnte.

Um den Abbau zu verhindern, wurde die Expressionsmethode folgendermaßen geändert: Nach Transformation des Vektors pET41b\_mod\_GST\_mut8 in den Expressionsstamm *E. coli* Rosetta (DE3) erfolgte eine Anzucht in 4 l M9-Minimalmedium mit zugesetztem  $^{15}$ N-markiertem Ammoniumchlorid auf eine OD<sub>600</sub> von 1,5 (statt 0,8 wie bisher). Die Expression von GST\_mut8 erfolgte über 5 h bei 28 °C nach Zugabe von 0,4 mM IPTG. Abbildung 3.20 zeigt Ausschnitte 12,5 %iger SDS-Gele nach Auftragen der gesammelten Fraktionen aus der GSH-Reinigung nach herkömmlicher Expression (links, vgl. auch Abb. 3.19) und nach optimierter Kurzexpression (rechts). Es ist deutlich zu sehen, dass der C-terminale Abbau zwar nicht gestoppt, jedoch reduziert werden konnte. Für die weiteren NMR-spektroskopischen Untersuchungen wurde daher diese Kurzexpressionsmethode zur Gewinnung von  $^{15}$ N-markiertem mut8 verwendet.



**Abbildung 3.20.:** 12,5 %ige SDS-Gele nach GSH-Aufreinigung nach herkömmlicher Expression (links, siehe auch Abb. 3.19) und optimierter Kurzexpression (rechts). Aufgetragen sind jeweils 5  $\mu$ l der gesammelten Fraktionen nach der Aufreinigung über eine GSH-Affinitätssäule (Gel links Spuren 1 bis 9 bzw. Gel rechts Spuren 1 bis 7) und in Spur M der Marker (NEB). Rot eingrahmt sind die Haupt- und Abbaubanden des GST\_mut8, deutlich zu erkennen ist, dass der Abbau nach optimierter Kurzexpression abnimmt.



**Abbildung 3.21.:** Gelfiltrationsprofil nach Trennung der GST und mut8 und das dazugehörige 15 %ige SDS-Gel. Von Volumen 75 bis 90 ml ist die Elution der GST zu sehen, von 105 bis 130 ml die Elution von mut8. In Spur M sind 4  $\mu$ l des Markers (Mark 12) aufgetragen. 1  $\mu$ l des einkonzentrierten GST\_mut8-Gemisches vor Zugabe der PreScission-Protease befindet sich in Spur 1. Vom einkonzentrierten GST\_mut8-Gemisch nach Zugabe und Inkubation der PreScission-Protease wurde 1  $\mu$ l in Spur 3 aufgetragen, danach folgen 2  $\mu$ l einer Fraktion (Spur 5) bei 85 ml (GST-Dimer). Je 10  $\mu$ l von 5 gesammelten Fraktionen wurden anschließend aufgetragen (Spuren 6 bis 10).

Nach Einkonzentrieren aller GST\_mut8-haltigen Fraktionen mithilfe eines Centricons (10.000 MWCO) wurde der GST-tag durch Zugabe von 28  $\mu$ g PreScission-Protease zu 12 mg GST\_mut8 von mut8 abgeschnitten. Im Anschluss daran wurde eine Gelfiltration (Superdex 75G 16/60 (GE-Healthcare)) durchgeführt, um beide Fragmente voneinander zu trennen.

Das Ergebnis der Gelfiltrationschromatographie ist in Abbildung 3.21 zu sehen. Das Gelfiltrationsprofil zeigt bei etwa 85 ml ein Signal (125 mAU), das der Fraktion des eluierten GST-Dimers zuzuordnen ist. Bei etwa 110 ml ist ein zweites Signal zu sehen (12 mAU), das dem eluierten mut8 (4,2 kDa) entspricht. In Abb. 2.4.2 ist neben dem Elutionsprofil das entsprechende 15%ige SDS-Gel eingefügt, in dem die erfolgreiche Trennung des Eluats sichtbar ist.

Nach der Gelfiltration wurden alle mut8-enthaltenen Fraktionen gesammelt und in einem 2000 MWCO (*Molecular Weight Cut Off*)-Centricon bis auf 600  $\mu$ l einkonzentriert. Während des Einkonzentrierens wurde das Peptid für die anschließenden NMR-Messungen in 50 mM KPi-Puffer, pH 6,5 umgepuffert.

## 3.4. Dynamik von mut8

Die Klonierung, Expression und Reinigung von mut8 bildeten den Ausgangspunkt zur Charakterisierung der dynamischen Eigenschaften des Peptids, durch welche sich ein vollständiges Bild bzgl. Struktur und Dynamik von mut8 ergeben sollte.

Das  $^{15}\text{N}$ -markierte Peptid lag jedoch nicht in Reinform vor, sondern war zusätzlich durch dessen Abbauprodukte verunreinigt. Daher musste vor der eigentlichen Dynamik-Messung eine Möglichkeit gefunden werden, die Signale der Abbauprodukte aus den folgenden NMR-Experimenten von denen des vollständigen mut8 zu unterscheiden, damit nicht fälschlicherweise ein Abbauprodukt in seiner Dynamik untersucht würde.

Um dies zu gewährleisten, wurde das nicht markierte, chemisch synthetisierte mut8 verwendet und über ein spezielles Messverfahren (SOFAST-HMQC, siehe Kap. 2.5.3) ein heteronukleares  $^1\text{H}$ - $^{15}\text{N}$ -Spektrum aufgenommen. Dieses macht sich die natürlich vorkommenden 0,3 % des  $^{15}\text{N}$ -Isotops zunutze. Das  $^{15}\text{N}$ -markiert vorliegende Peptid wurde anschließend dazu verwendet, ein heteronukleares HSQCNOE-Spektrum aufzunehmen.

Dynamische Prozesse sind dabei abhängig von der Größe eines Proteins bzw. Peptids; je kleiner ein Peptid ist, desto schneller bewegt es sich. Villin Headpiece 36 (1vii) faltet sich zu seiner finalen Tertiärstruktur innerhalb von etwa 4,3  $\mu\text{s}$  [104] und gilt damit als ein sich sehr schnell faltendes Peptid. Eine interne Dynamik, die dazu führt, dass mut8 keine stabile Tertiärstruktur ausbilden kann, würde dagegen im ps- bis ns-Bereich zu erwarten sein. HSQCNOE-Spektren eignen sich besonders gut für die Charakterisierung der Dynamik von Peptiden und Proteinen im ps- bis ns-Bereich.

### 3.4.1. Aufnahme des SOFAST-HMQC und Zuordnung

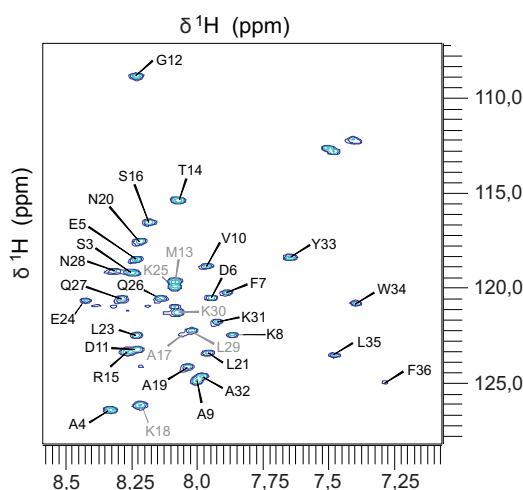
Mit der bereits für die Aufnahme der COSY-, TOCSY- und NOESY-Spektren verwendeten Probe des unmarkierten synthetischen mut8 wurde ein SOFAST-HMQC aufgenommen [86]. In diesem Spektrum werden die Kopplungen zwischen den N und den daran gebundenen  $\text{H}_\text{N}$  als Signale sichtbar.

Das Spektrum wurde mit 2000 scans und somit über 20 h Messzeit aufgenommen, da das Isotop  $^{15}\text{N}$  eine natürliche Häufigkeit von nur 0,3 % besitzt. Bei einer  $^{15}\text{N}$ -markierten Probe würde die Messzeit im Vergleich nur etwa 7 bis 40 min dauern.

Im SOFAST-HMQC erhält man für jede  $\text{H}_\text{N}$ -N-Kopplung einer Aminosäure ein Signal, da jede Aminosäure (außer Prolin) ein Rückgrat-NH besitzt. Zusätzlich werden die Signale sichtbar, welche durch aminogruppenhaltige Seitenketten der Aminosäuren wie Glutamin und Asparagin zustande kommen.

Abbildung 3.22 zeigt den für die Zuordnung und die nachfolgenden Messungen relevanten Ausschnitt aus dem aufgenommenen SOFAST-HMQC inklusive der Zuordnung der Aminosäuren. Dabei sind die Aminosäuren M13, A17, K18, K25, L29 und K30 grau gekennzeichnet, weil sie nicht eindeutig zugeordnet werden konnten.

Für die Zuordnung wurden die  $\text{H}_\text{N}$ -Werte aus dem TOCSY-Spektrum verwendet und die  $\text{N}_\text{H}$ -Werte nach Wishart *et al.*, 1995 (Tabelle 3) [90] hinzugezogen.



**Abbildung 3.22.:** SOFAST-HMQC mit Zuordnung. Nicht zugeordnet sind Aminosäuren M1 und Y2, nicht eindeutig bestimmte Aminosäuren und deren Zuordnungspfeile sind in grau dargestellt.

Zusätzlich dazu wurden die  $N_H$ - $H_N$ -Werte aus der Zuordnung des Wildtyps von Villin Headpice 36 (1vii) von McKnight *et al.*, 1996 [16] zur eindeutigen Zuordnung hinzugezogen. Diese Werte basieren auf einer Sequenz, die sich um 8 Aminosäuren von der hier vorliegenden mut8-Sequenz unterscheidet (da mut8 die Achtfachmutante des 1vii darstellt) und bei 30 °C und einem pH von 5 aufgenommen wurden (vgl. hier: 25 °C und pH 6,5). Daher konnten die Zuordnungen nur für V10 und G12 übernommen werden.

M1, Y2 und P22 konnten ebenfalls nicht bestimmt werden, da für M1 und Y2 keine  $H_N$ -Zuordnungen existieren und Prolin 22 aufgrund des fehlenden  $H_N$  generell im SOFASTHMQC-Spektrum nicht sichtbar ist. Für jede Kopplung war im Spektrum genau ein Signal vorhanden, daher ist es unwahrscheinlich, dass multiple stabile Konformere vorliegen bzw. ein langsamer Austausch zwischen zwei Formen von mut8 vorliegt.

Die Schwierigkeiten bei der Berechnung einer Tertiärstruktur könnten also an dem schnellen dynamischen Verhalten des Peptids liegen. Der Nachweis dieser Dynamik von mut8 kann über die Bestimmung des hetNOE erfolgen (siehe Kap. 3.4.2).

### 3.4.2. Aufnahme des HSQCNOE und Berechnung des hetNOE

Im Anschluss an die Aufnahme und Zuordnung des SOFAST-HMQC wurde ein HSQCNOE aufgenommen [87, 88]. Dazu wurde eine 200  $\mu$ M Probe des  $^{15}$ N-markierten mut8 in 600  $\mu$ l KPi-Puffer (50 mM), pH 6,5, verwendet (siehe Kap. 2.5.1). Für die Aufnahme des HSQCNOE wurde das Pulsprogramm hsqcnoef3gpsi der Pulsprogramm-Bibliothek von Bruker verwendet. Die Pulsfolge sorgt für eine alternierende gleichzei-

tige Aufnahme zweier HSQCNOE-Spektren, bei denen eines mit Vorsättigung (NOE-Aufbau) und eines ohne Saturierung aufgenommen wird.

Das resultierende Spektrum wurde nach Aufnahme mit Hilfe des Bruker-Programms Topspin 2.0 in seine beiden Teilspektren aufgespalten. Die beiden resultierenden HSQCs sind jedoch keine gewöhnlichen HSQC-Spektren, sondern welche, dessen Signale durch Kopplungen zwischen dem  $H_N$  und  $N_H$  über den Raum – nicht über Bindungen – zustande kommen. Nach Fouriertransformation der aufgenommenen Datenpunkte erhält man für das Spektrum ohne Saturierung ausschließlich positive (oder negative) Signale und für das Spektrum mit Saturierung sowohl positive als auch negative Signale.

Aus dem ersten Spektrum lassen sich die Signal-Intensitäten  $I_{eq}$  (Intensitätgleichgewicht) berechnen, aus dem zweiten Spektrum die Signal-Intensitäten  $I_{sat}$  (Intensitätsaturierung).

Die maximale NOE-Signalverstärkung  $\eta_{max}$  wird über den Quotienten aus den gyromagnetischen Verhältnissen von  $\gamma_S$  zu  $\gamma_I$  bestimmt (siehe Gleichung 3.1), dabei steht  $\gamma_S$  für den Nukleus, der saturiert wird (hier:  $^1H$ ), und  $\gamma_I$  für den Nukleus, dessen Verstärkung gemessen wird (hier:  $^{15}N$ ) [105].

$$\eta_{max} = \gamma_S / 2\gamma_I \quad (3.1)$$

Dadurch, dass das gyromagnetische Verhältnis  $\gamma$  über Gleichung 3.2 (mit  $|B|$ =Stärke des Magnetfeldes) mit der Larmorfrequenz  $\omega$  korreliert ist, lässt sich  $\eta_{max}$  auch in Abhängigkeit von der Larmorfrequenz des untersuchten Nukleus  $\omega_x$  und der Rotationskorrelationszeit  $\tau_c$  darstellen. Letztere beschreibt die Rotationsbewegung des gesamten Proteins und ist proportional zur Molekülgröße.

Je größer also das zu untersuchende Peptid oder Protein, desto größer ist auch  $\tau_c$ . Ist  $\omega_N \cdot \tau_c \gg 1$ , so handelt es sich um einen langsam rotierenden Grenzfall, ist  $\omega_N \cdot \tau_c \ll 1$ , so wird ein schnell rotierender Grenzfall beschrieben [106].

$$\omega = (\gamma / 2\pi) \cdot |B| \quad (3.2)$$

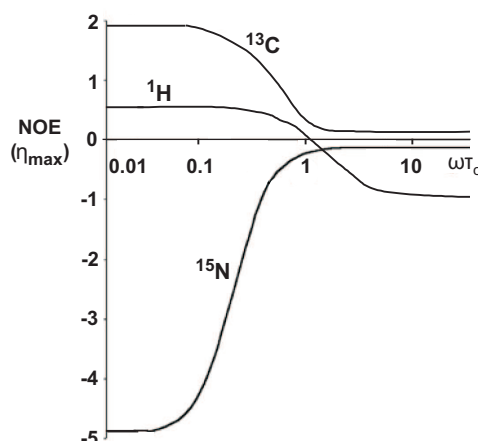
Abbildung 3.23 zeigt die Abhängigkeit von  $\eta_{max}$  zu  $\omega_x \cdot \tau_c$  für  $^1H$ ,  $^{13}C$  und  $^{15}N$ . Für den Fall, dass der Quotient aus  $\gamma_S$  und  $\gamma_I$  der jeweiligen Kerne negativ ist, ist auch  $\omega$  negativ.

Für den Quotienten aus  $\gamma_S / 10^7 = 26,75$  (für  $^1H$ ) und  $\gamma_I / 10^7 = -2,71$  (für  $^{15}N$ ) ist dies nach Einsetzen in Gleichung 3.1 mit

$$\gamma_{^1H} / 2 \cdot \gamma_{^{15}N} = -4,94$$

der Fall ( $[\gamma_x] = \text{radT}^{-1}\text{s}^{-1}$ ). Da es sich bei mut8 um ein kleines Peptid handelt und es sich wahrscheinlich schnell bewegt, lässt sich schlussfolgern, dass  $\omega_N \cdot \tau_c < 1$  ist, man somit NOE-Signalverstärkungen im Bereich  $< -0,5$  erwartet würde.

Nun lässt sich, nach Interpretation und Auswertung der Spektren, eine Gesamtbewegung bestehend aus den Fluktuationen der  $H_N$ s, der Bewegungen der Sekundärstrukturelemente wie  $\alpha$ -Helices und des ganzen Proteins bzw. Peptids berechnen. Dieser sogenannte hetNOE (heteronuklearer NOE) ( $\eta$ ) wird berechnet, indem der Quotient



**Abbildung 3.23.:** Abhängigkeit der maximalen theoretischen NOE-Intensität von  $\omega_x \cdot \tau_c$  für X( $^1\text{H}$ )-Experimente; X= $^1\text{H}$  (homonuklearer NOE),  $^{13}\text{C}$  und  $^{15}\text{N}$ . Der Null-durchgang für homonukleare NOEs liegt bei  $\omega_x \cdot \tau_c = 1,12$  [107]. Verändert nach [105, 107].

aus  $I_{sat}$  und  $I_{eq}$  gebildet und -1 genommen wird (siehe Gleichung 3.3). Der hetNOE kann laut Definition Werte zwischen -3,94 ( $\eta_{max}$ ) und 0,78 annehmen.

$$\eta = (I_{sat}/I_{eq}) - 1 \quad (3.3)$$

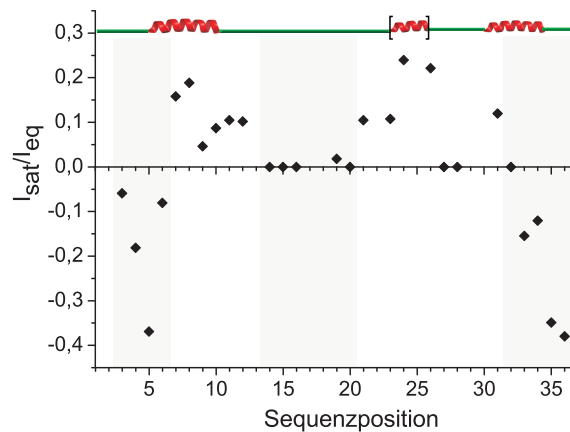
In der Literatur wird der hetNOE jedoch meist nur als Quotient aus  $I_{sat}/I_{eq}$  dargestellt [105]. Auch in dieser Arbeit wird darauf verzichtet, die Subtraktion von -1 durchzuführen, da so die Werte besser mit denen der Literatur verglichen werden können.

Darin ist zu finden, dass der hetNOE Werte zwischen -1 und +1 annehmen kann. Wird ein Protein als starr definiert, so liegt der theoretische Wert bei +1, liegt der Wert bei  $>+0,6$ , so kann man davon ausgehen, dass das Protein sehr wenig flexibel ist. Bei sehr großer Flexibilität werden negative Werte erwartet [108].

Nach Aufnahme des HSQCNOE wurden mit Hilfe der bereits zugeordneten Signale im SOFAST-HMQC die Signale im HSQCNOE zugeordnet. Im Anschluss daran wurden mit Hilfe des Programms Sparky [79] über eine Gauß-Funktion die Volumina der einzelnen Resonanzen ausgelesen und über Formel 3.3 der hetNOE berechnet. Als Kontrolle wurden im Programm Auremol die Intensitäten aller Signale berechnet und ebenfalls der Quotient  $I_{sat}/I_{eq}$  gebildet. Da die Ergebnisse aus beiden Rechnungen nahezu identisch waren, konnte von einer korrekten Berechnung ausgegangen werden. Abbildung 3.24 zeigt die berechneten hetNOE-Werte für alle Aminosäuren von mut8. Die unsicher zugeordneten Aminosäuren M13, A17, K18, K25, L29 und K30 wurden zur Berechnung nicht verwendet, daher befinden sich an diesen Stellen Lücken in der Darstellung.

Für den Quotienten aus  $I_{sat}$  und  $I_{eq}$  gilt, dass ein negativer Wert für einen flexiblen





**Abbildung 3.24.:** Darstellung der Berechnung von  $I_{sat}/I_{eq}$ . Je negativer der Quotient, desto flexibler sind die Bereiche von mut8; grau unterlegt sind besagte Regionen und reichen von Aminosäuren 3 bis 6, 14 bis 20 und 32 bis 36.

Bereich spricht, während positive Werte für stabile Regionen innerhalb eines Peptides sprechen. Je negativer der Wert ist, desto flexibler ist der jeweilige Bereich. Wie man Abbildung 3.24 entnehmen kann, handelt es sich bei den grau unterlegten Bereichen um die Regionen in mut8, welche ein erhöhtes dynamisches Verhalten aufweisen. Dabei ist zu erkennen, dass sich die Regionen großer Flexibilität zwischen Aminosäuren 3 und 6 und zwischen 32 und 36 befinden, ein etwas weniger flexibler Bereich befindet sich zwischen Aminosäuren 14 und 20.

## 4. Diskussion

Bei Villin Headpiece 36 (pdb-ID: 1vii) handelt es sich um ein stabil faltendes Peptid aus 36 Aminosäuren. Wird Aminosäure F18 gegen ein Lysin ausgetauscht, so wird aus dem stabilen Peptid eine strukturell instabile Mutante. Diese sollte über den evolutionären Algorithmus EMOS wieder in die stabile Struktur zurückgeführt oder sogar strukturell noch stabiler als der Wildtyp werden. Dabei sollte gleichzeitig eine Liganden-erreichbare Schleife vorhanden sein (vgl. Kap. 1).

Aus dem Resultat dieser Vorhersage über EMOS, einem Cluster aus mehreren potentiellen Sequenzen, wurde die Mutante r1g11i1 ausgewählt. Diese besaß acht Mutationen gegenüber dem Wildtyp und sollte oben genannte Kriterien (strukturelle Stabilität, Vorhandensein einer Liganden-Schleife) erfüllen. Für die Sequenz von r1g11i1 wurde eine Struktur vorhergesagt: new2. Diese besitzt, wie auch HP36 (wt), drei  $\alpha$ -Helices und sollte in einer stabilen Konformation vorliegen.

Um nun vergleichen zu können, ob EMOS' Vorhersage bzgl. r1g11i1 korrekt war, wurde über NMR-Spektroskopie die Struktur des chemisch synthetisierten mut8 näher charakterisiert, um diese dann mit new2 zu vergleichen.

Die NMR-spektroskopischen Auswertungen dieser Arbeit ergaben ein erstaunliches Resultat. So konnten zwar die Sekundärstrukturelemente gut beschrieben und definiert werden (siehe Kap. 3.1.2), jedoch konnte für das Peptid keine eindeutige Tertiärstruktur bestimmt werden. Vielmehr wurde entdeckt, dass es sich bei mut8 um ein dynamisches Peptid handelt, welches im ps- bis ns-Bereich verschiedene Konformationen einnimmt. Das Kriterium der strukturellen Stabilität wurde somit – laut NMR-spektroskopischer Ergebnisse – nicht erfüllt.

Unter der Annahme, dass die *in vitro* erhobenen Daten korrekt sind, müsste auch eine *in silico*-Analyse von new2 dieselben Ergebnisse liefern. Passen beide Resultate zueinander, dann lässt sich erstens aus den so gewonnenen Erkenntnissen EMOS verbessern und zweitens eine Auswertemethode entwickeln, mit der schnell und effizient nach einem erneuten EMOS-Durchlauf potentielle Sequenzen als gute oder weniger gute Kandidaten selektiert werden können.

Neben den NMR-spektroskopischen Analysen von mut8 wurden *in silico*-Untersuchungen durchgeführt, die sich mit der Vorhersage der Sekundärstrukturelemente über DSSP (*Define Secondary Structure of Proteins*) und NOE-Berechnungen beschäftigten. Ebenfalls wurde analog zu der Winkelvorhersage über TALOS eine bioinformatische Diederwinkelvorhersage erstellt und mit diesen neuen Parametern ein weiteres Strukturensamble berechnet sowie RMSF (*Root Mean Square Fluctuation*)-Kalkulationen durchgeführt, die einen Vergleich mit der über NMR bestimmten Dynamik von mut8 ermöglichen.

Das nun folgende Kapitel analysiert und vergleicht die zu mut8 erhobenen experimentellen Ergebnisse zur Sekundärstruktur mit den bioinformatischen Simulationen,

welche mit new2 durchgeführt wurden. Danach erfolgt ein Vergleich der über TALOS ermittelten  $\phi$ - und  $\psi$ -Werte mit den Ergebnissen aus der Bioinformatik. Im Anschluss daran wird die Tertiärstrukturberechnung von mut8 diskutiert, um dann auf die Expression und Reinigung einzugehen. Abschließend werden die über NMR-Spektroskopie gewonnenen Ergebnisse zur Dynamik von mut8 mit denen aus der Bioinformatik verglichen. Alle hierbei verwendeten Bioinformatik-Daten wurden, wenn nicht anders erwähnt, von Manuel Prinz (AG Prof. Daniel Hoffmann, ZMB, Universität Duisburg-Essen) erhoben.

## 4.1. mut8 hat eine gut definierte Sekundärstruktur

Von dem chemisch synthetisierten Peptid wurde zunächst durch Aufnahme eines 1D-NMR-Spektrums festgestellt, dass wahrscheinlich weder mehrere Populationen von mut8 existieren noch verschiedene stabile Konformationen gleichzeitig vorliegen. Somit konnte das Produkt der Synthese als rein bezeichnet werden.

Weiterhin konnte man bereits aufgrund der hohen Dispersion der Frequenzen und der Verschiebungen der  $H_N$ -Signale nach niederen ppm-Werten erkennen, dass es sich bei mut8 um ein strukturiertes Peptid handelt und eher  $\alpha$ -helikale Anteile besitzt (siehe Kap. 3.1.1).

Daraufhin wurden homonukleare 2D-Spektren (COSY, TOCSY, NOESY) aufgenommen. Die Zuordnung der Spektren erfolgte über das *Chaintracing*-Verfahren; mut8 stellte mit 36 Aminosäuren ein kleines Peptid dar und zeigte somit eine übersichtliche Anzahl von Signalen in den Spektren. Zudem wurde die Zuordnung dadurch erleichtert, dass die meisten Resonanzen gut im Spektrum verteilt waren. Die Analyse der homonuklearen Spektren reichte aus, um die Sekundärstruktur gut zu charakterisieren (siehe Kap. 3.1.2).

### 4.1.1. NOEs und CSI – Resultate aus der NMR-Spektroskopie

Nach der Zuordnung der Spektren (siehe Kap. 3.1.1) wurden die Abstände einzelner Protonen voneinander ermittelt, die  $\leq 5 \text{ \AA}$  waren (Konnektivitäten über NOEs) und über welche direkte Aussagen über die Sekundärstrukturverteilung getroffen werden können (siehe Kap. 3.1.2).

Von den 59 gefundenen NOEs waren 12 (20 %) nicht genau bestimmbar, da sie aufgrund von Signalüberlagerungen im NOESY-Spektrum nicht eindeutig zuzuordnen waren.

Ausschlaggebend für das Vorhandensein von  $\alpha$ -Helices ist laut Wagner *et al.*, 1986 [92], dass zumeist  $d_{\alpha\beta}(i, i+3)$ - sowie  $d_{\alpha N}(i, i+3)$ -,  $d_{\alpha N}(i, i+4)$ - und  $d_{NN}(i, i+2)$ -Konnektivitäten zu sehen sind. Dies war bei mut8 jedoch nur begrenzt der Fall (siehe Kap. 3.1.2). So fehlen viele  $d_{\alpha\beta}(i, i+3)$ -Konnektivitäten, die am aussagekräftigsten sind, um eindeutige  $\alpha$ -Helix-Bereiche in mut8 festzulegen.

Stattdessen finden sich 26  $d_{\alpha N}(i, i+1)$ -Konnektivitäten, welche laut Wüthrich 1986 [91] eher einem  $\beta$ -Faltblatt-Bereich zuzuordnen sind. Jedoch erwähnt er auch gleichzeitig, dass die Eindeutigkeit der Bestimmung einer  $\beta$ -Faltblatt-Region über  $d_{\alpha N}(i,$

$i+1$ )-Konnektivitäten wesentlich uneindeutiger ist als die eines Helix-Bereiches anhand von auftretenden  $d_{NN}(i, i+1)$ -Konnektivitäten. Da insgesamt 14 dieser NOEs (davon nur 3 ungenau) für mut8 bestimmt werden konnten, konnten mit Hilfe dieser und der übrigen  $\alpha$ -Helix-spezifischen NOEs ungefähre Helixbereiche definiert werden. So wurden  $\alpha$ -Helices im Bereich von Aminosäuren 5 bis 11 und 29 bis 35 (relativ gut definiert) sowie von Aminosäuren 22 bis 25 (weniger gut definiert) bestimmt.

Zusätzlich zu dieser Form der Sekundärstrukturbestimmung wurde auch der *Chemical Shift Index* (CSI) berechnet (siehe Kap. 3.1.2). Da für diese Methode die  $H_\alpha$ -Resonanzen benötigt werden und diese zu 100 % zugeordnet werden konnten, ergab sich hier ein präziseres Bild der Sekundärstrukturverteilung als nach der Auswertung der Konnektivitäten über NOEs. So konnten  $\alpha$ -Helices von Aminosäuren 4 bis 8, 30 bis 33 und 23 bis 26 bestimmt werden, wobei die letztgenannte Region nicht genau definiert ist, da die  $+1$ -Werte in diesem Bereich durch einen 0-Wert unterbrochen werden (vgl. Kap. 3.12). Somit entspricht dieser Bereich nicht mehr exakt der Definition einer  $\alpha$ -Helix nach Wishart *et al.*, 1991. Ein Vergleich der über NOEs und dem CSI erhobenen Werte zeigt, dass sich diese sehr ähnlich sind; so ergeben sich folgende  $\alpha$ -helikale Bereiche als arithmetische Mittelung der NOE- und der CSI-Daten: Aminosäuren 5 bis 10, 30 bis 35 und 23 bis 26. Interessanterweise zeigen sowohl die Ergebnisse der Konnektivitäten über NOEs als auch die des CSI, dass die mittlere  $\alpha$ -Helix zwischen Aminosäuren 23 und 26 weniger gut definiert ist.

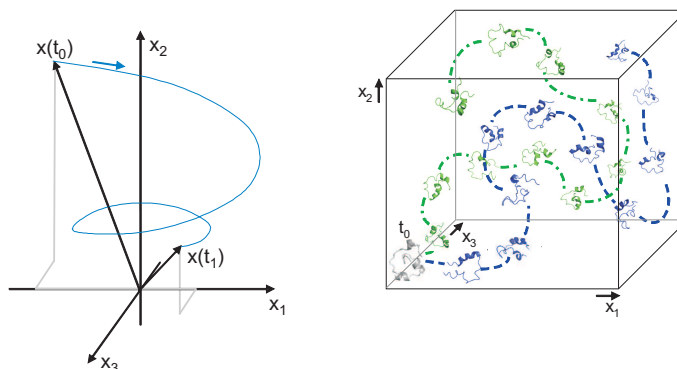
#### 4.1.2. NOEs und DSSP – Resultate aus der Bioinformatik

Analog zu der experimentellen Bestimmung der NOEs anhand der 2D-NMR-Spektren wurden von Manuel Prinz auch über bioinformatische Methoden Abstände zwischen Protonen über vorhandene NOEs berechnet.

Die zu Beginn des Kapitels erwähnte Struktur new2, die sich laut Vorhersage für die Sequenz von r1g11i1 ergeben sollte, wurde als Startstruktur für Molekulardynamik-Simulationen (MD-Paket GROMACS, [www.gromacs.org](http://www.gromacs.org)) gewählt. Ausgehend von new2 wurden zwei MD-Simulationen über 10 ns durchgeführt (zwei Experimente). Die beiden Experimente unterscheiden sich dadurch, dass der Startstruktur new2 unterschiedliche Sätze an Geschwindigkeitsvektoren zugeordnet werden. Aus jedem Experiment wurden 64 Trajektorien für 64 verschiedene Temperaturen generiert. Aus Experiment 1 und 2 wurde dann jeweils die Trajektorie bei 300 K zur weiteren Auswertung ausgewählt.

Als Trajektorie bezeichnet man eine Serie von Momentaufnahmen einer Struktur, die sich z. B. in einem temperierten Wasserbad in einem definierten, virtuellen Raum bewegt [9]. In den Trajektorien nimmt die Ausgangsstruktur erstens verschiedene Konformationen ein und geht zweitens einen völlig willkürlichen Weg durch diesen Raum.

Anhand einer einzelnen Startstruktur können sehr viele Trajektorien generiert werden, da sich diese immer rein zufällig durch den Raum bewegt und sich somit immer andere Momentaufnahmen ergeben. Jede Trajektorie besteht hier aus 1000 Momentaufnahmen der Struktur. Schematisch ist das Prinzip einer Trajektorie auf Vektorebene im 3D-Raum und die Vorstellung zweier solcher Trajektorien in Abbildung 4.1 gezeigt. Darin bildet die graue Struktur, welche in Experiment 1 Geschwindigkeitsvektor  $a$  und



**Abbildung 4.1.:** Schematische Zeichnung einer Trajektorie in Vektor-Darstellung (links) und Übertragung davon auf die zwei möglichen Trajektorien eines Peptids in einem Raum (gefüllt mit Wasser, bei Raumtemperatur) ausgehend von einer Startstruktur (grau). Diese wählt einmal den blauen Weg mit den blauen Konformationen durch den Raum und einmal den grünen Weg mit den grünen Konformationen (rechts).

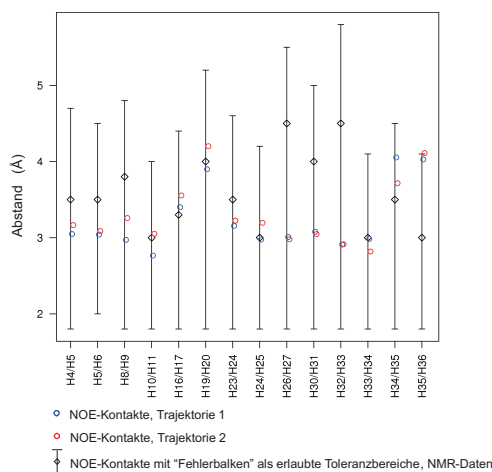
in Experiment 2 Geschwindigkeitsvektor  $b$  zugeordnet bekommt, den Ausgangspunkt und geht zufällig einmal den blauen Weg durch den Raum mit den unterschiedlichen blauen Strukturen (Experiment 1, Trajektorie 1 bei 300 K, 10 Momentaufnahmen) und einmal den grünen Weg mit den verschiedenen grünen Strukturen (Experiment 2, Trajektorie 2 bei 300 K, 10 Momentaufnahmen).

Die Startstruktur war für den hier beschriebenen Fall new2. Für new2 wurden besagte zwei Trajektorien aus je 1000 Strukturen generiert und diese unterschiedlich ausgewertet. Der erste Auswertungsschritt beinhaltete die Ermittlung der NOEs.

Aus den Trajektorien, die jeweils aus 1000 möglichen Strukturen besteht, wurden alle sichtbaren NOEs der Art  $d_{\alpha N}(i, i+1)$ ,  $d_{NN}(i, i+1)$ ,  $d_{NN}(i, i+2)$ ,  $d_{\alpha N}(i, i+2)$ ,  $d_{\alpha N}(i, i+3)$ ,  $d_{\alpha N}(i, i+4)$  und  $d_{\alpha\beta}(i, i+3)$  bestimmt.

Exemplarisch sind die Ergebnisse der bioinformatischen NOE-Berechnungen aus zwei Trajektorien für  $d_{NN}(i, i+1)$  in Abbildung 4.2 dargestellt, alle übrigen Ergebnisse befinden sich im Anhang. Dazu wurden die Abstände zwischen zwei Atomen (hier die  $H_{Ns}$  aller Atome) von den 1000 Strukturen aus den zwei Trajektorien gemessen und jeweils die aus einer Trajektorie (siehe Abb. 4.2, blaue bzw. rote Kreise) für jeden gefundenen NOE gemittelt.

Zusätzlich dazu sind die Abstände zwischen im NMR-Spektrum existierenden NOEs der Sorte  $d_{NN}(i, i+1)$  abgebildet, die Rauten stehen dabei für den berechneten Abstand  $d$  zwischen den zwei Atomen. Die Fehlerbalken stellen den Toleranzbereich dar. Dieser ergibt sich aus einem Mindestabstand zwischen zwei Atomen von  $1,8 \text{ \AA}$  und einem berechneten Maximalabstand von  $d+x \text{ \AA}$ . Alle bioinformatisch erhobenen Daten, sowohl aus Trajektorie 1 als auch aus Trajektorie 2, liegen im Toleranzbereich der aus NMR-Messungen bestimmten Daten. Nach Vergleich aller erhobenen Distanzen, welche über NOEs ermittelt wurden (siehe Kap. A), liegen insgesamt 76 % der mit Hilfe von bioinformatischen Methoden bestimmten Distanzen auf den aus NMR-Messungen berechneten Werten bzw. in deren Toleranzbereich.



**Abbildung 4.2.:** Ein Beispiel der Distanz-Bestimmungen  $d_{NN}(i, i+1)$  über NOEs aus den zwei Trajektorien (blaue und rote Kreise) im Vergleich zu den Distanzen aus den NMR-Messungen (Rauten). Die Fehlerbalken dienen hier als Anzeige für die nach Auremol erlaubten Toleranzbereiche der Abstände zwischen zwei Atomen (min. 1,8 Å).

Die NMR-spektroskopisch ermittelten Distanzen basieren auf drei Auswerteparametern: 1. Der NOE muss im NOESY-Spektrum vorhanden sein. 2. Die Intensität des NOEs wird ausgewertet und in schwach, mittel oder stark eingestuft. 3. Es wird geprüft, ob der NOE eindeutig oder uneindeutig ist.

Die Intensitäten der NOEs liefern zusätzlich Informationen darüber, wie weit die zwei am NOE beteiligten Atome voneinander entfernt sind, so gilt für die NOE-Intensität mit  $r$  als Abstand zwischen zwei Protonen und  $\tau_c$  als Korrelationszeit (siehe Gleichung 4.1):

$$NOE \propto 1/r^6 \cdot f(\tau_c) \quad (4.1)$$

Je größer also der Abstand zwischen zwei Protonen ist, desto kleiner wird der NOE, woraus ein schwaches Signal im NOESY-Spektrum resultiert.

Die Bioinformatik verwendet nur einen Parameter zur Auswertung der Strukturen aus Trajektorie 1 und 2, der von den experimentell erhobenen Daten ausgeht: Wie groß ist der Abstand zwischen zwei koppelnden Atomen, welche am NOE beteiligt sind und dessen NOE gleichzeitig im NOESY-Spektrum zu sehen ist?

Somit können die bioinformatisch erhobenen Daten mit den experimentellen NMR-Daten nur über die Abstände zwischen zwei Protonen miteinander verglichen werden. Dies ist wiederum möglich, da über Gleichung 4.1 die Abstände der Protonen proportional zur bereits ermittelten Intensität des NOEs ist.

Nach Vergleich aller erhobenen Distanzen über NOEs (siehe Anhang B) liegen insgesamt 77 % der vorhergesagten Atomabstände im Toleranzbereich der experimentellen Daten, was für eine gute Übereinstimmung spricht. Die von den NMR-Daten abwei-

chenden Ergebnisse finden sich hauptsächlich im Bereich zwischen Aminosäuren 13 und 19. So liegen die Distanzen für  $d_{\alpha\beta}$  ( $i, i+3$ ) von Aminosäure 14 bis 17 und 16 bis 19 sowie die Distanzen für  $d_{\alpha N}$  ( $i, i+4$ ) von Aminosäure 13 bis 17 nicht mehr im Toleranzbereich der NMR-Daten. Ebenfalls abweichend verhalten sich die Werte für die Distanzen  $d_{NN}$  ( $i, i+2$ ) von Aminosäure 9 bis 11, 23 bis 25, 29 bis 31 und 33 bis 35.

Zusätzlich zu der Bestimmung der Distanzen über NOEs *in silico* wurde eine zweite Methode verwendet, um die insgesamt 2000 Strukturen der beiden Trajektorien auf den Sekundärstrukturanteil hin zu untersuchen. Diese Methode nennt sich DSSP (*Define Secondary Structure of Proteins*) und wurde von Kabsch und Sander bereits 1983 entwickelt [109].

DSSP startet mit der Erkennung von Wasserstoffbrücken innerhalb eines Proteins bzw. Peptids und fügt geometrische Informationen aus Röntgenkristallographie-Strukturen hinzu. Sekundärstrukturen werden dann als Wiederholungen von elementaren H-Brücken-Mustern in „Schleifen“ und „Brücken“ eingeteilt. Sich wiederholende „Schleifen“ sind Helices, sich wiederholende „Brücken“ sind „Stufen“, verbundene „Stufen“ wiederum werden als  $\beta$ -Faltblätter definiert.

Die Lösungsmittel-Exposition wird dabei als Anzahl der Wassermoleküle im möglichen Kontakt mit einer Aminosäure gegeben [109]. DSSP ermittelt so insgesamt 8 verschiedene Kategorien von Sekundärstrukturelementen: Die  $3_{10}$ -Helix,  $\alpha$ - und  $\pi$ -Helices (anhand von sich wiederholenden Wasserstoffbrücken-Sequenzen zwischen Aminosäuren  $i$  und  $i+3$ ,  $i+4$  bzw.  $i+5$ ),  $\beta$ -Faltblätter (antiparallel, parallel und einzelne Stränge), Schleifen und andere.

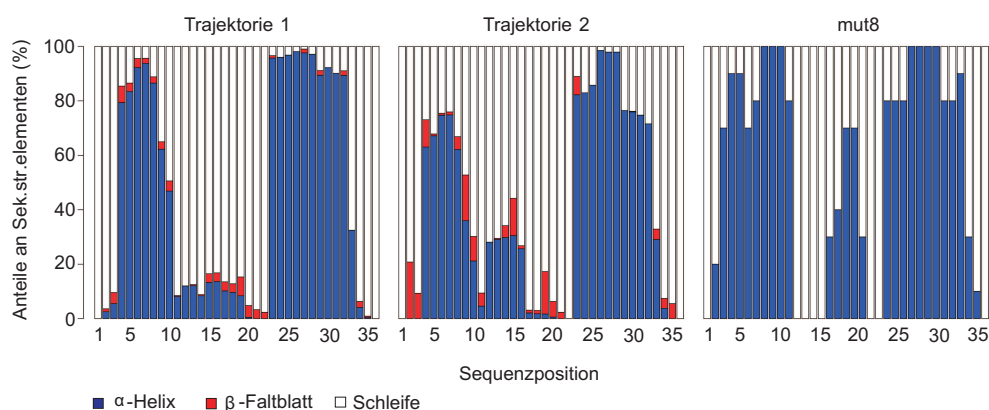
Mit Hilfe der DSSP wurden so von Manuel Prinz jeweils die prozentualen Anteile der Sekundärstrukturelemente über das gesamte Ensemble der Trajektorien 1 und 2 ermittelt und als Balkendiagramm dargestellt.

Dieselbe Methode wurde auch auf das Ensemble von mut8 (siehe Kap. 3.1.3) angewendet. In Abbildung 4.3 sind die Balkendiagramme für die Resultate beider Trajektorien (links, Mitte) und für das mut8-Ensemble (rechts) dargestellt. Die blauen Balken stellen dabei die prozentualen Anteile für  $\alpha$ -helikale Bereiche dar, rote Balken definieren mögliche  $\beta$ -Faltblatt-Regionen und weiße Balken Schleifen-Bereiche.

Über DSSP wurden aus Trajektorie 1 und 2  $\alpha$ -helikale Bereiche zwischen den Aminosäuren 4 und 10 sowie 23 und 32 ermittelt. Die Ergebnisse für Trajektorie 2 zeigen, dass die erste Helix (Aminosäuren 4 bis 10) mit etwa 60 % Auftrittswahrscheinlichkeit schlechter definiert ist als die entsprechende Helix bei Trajektorie 1 (80 %).

Gleichzeitig taucht in der zweiten Trajektorie mit einem Anteil von etwa 30 % eine Helix zwischen den Aminosäuren 12 bis 16 auf. Die DSSP-Berechnung für das berechnete Ensemble von mut8 zeigt eine N-terminale Helix von Aminosäuren 3 bis 11 (Anteil von 80 %) und eine C-terminale Helix mit einem prozentualen Anteil von ebenfalls 80 %. Auch in dieser Rechnung taucht eine mittlere  $\alpha$ -Helix zwischen Aminosäuren 16 und 20 mit einem Anteil von nur 40 % auf.

Helixbereiche der N- und C-terminalen Helices von mut8 stimmen mit den *in silico*-Werten gut überein. So ist die N-terminale Helix mit 80 % in Trajektorie 1, mit 60 % in Trajektorie 2 und mit 80 % in der mut8-Trajektorie definiert. Die C-terminale Helix ist in Trajektorie 1 mit 90 %, in Trajektorie 2 mit etwa 80 % und in der mut8-Trajektorie



**Abbildung 4.3.:** DSSP-Berechnungen für je 1000 Strukturen zweier Trajektorien von new2 (links, Mitte) und des Ensembles von mut8 (rechts). Die blauen Balken stellen die prozentualen Anteile von  $\alpha$ -Helices des jeweiligen Sequenzbereiches dar, die roten Balken die Anteile für das Auftreten von  $\beta$ -Faltblättern und die weißen Balken die für das Vorhandensein von Schleifen.

mit 90 % definiert. Bei Trajektorie 2 zeigt sich eine mittlere Helix (Anteil von 30 %), die bei mut8 noch deutlicher wird, jedoch sind die Begrenzungen (Aminosäuren 12 bis 16 für new2 bzw. 16 bis 20 für mut8) deutlich verschoben, sodass hier eine schlecht definierte mittlere Helix zu finden ist.

Eine Zusammenfassung, Gegenüberstellung und Mittelung (nur DSSP-Werte) der Sekundärstrukturverteilung in new2/mut8 lässt sich Tabelle 4.1 entnehmen. Darin enthalten sind die für mut8 bestimmten Sekundärstrukturverteilungen, die über die NOEs ermittelt wurden, sowie die über den CSI und die DSSP-Berechnung ermittelten Sekundärstrukturverteilungen.

**Tabelle 4.1.:** Auflistung der für new2/mut8 aufgrund der NOEs, des CSI und der DSSP ermittelten  $\alpha$ -helikalen Bereiche.

	NOEs (mut8)	CSI (mut8)	DSSP (new2)	DSSP (mut8)	$x_{arithm.}$	NOEs (pdb-ID:1vii)
Aminosäuren	5-11	4-8	4-10	3-11	4-11	4-11
	[22-25]	[23-26]	[12-16]	[16-20]	[14-18]	15-20
	29-35	30-33	23-32	23-32	23-32	23-34

Ebenfalls aufgeführt ist das DSSP-Ergebnis für new2. Von den Werten der DSSP-Rechnungen für new2 und mut8 wurde ein arithmetisches Mittel berechnet ( $x_{arithm.}$ ). Hier wird deutlich, dass die *in silico*-Werte mit den *in vitro*-Daten übereinstimmen. Die einzige signifikante Abweichung findet sich für die mittlere  $\alpha$ -Helix, die einmal im Bereich von Aminosäuren 22 bis 25 bzw. 23 bis 26 definiert ist (NOEs und CSI), während sie bei den DSSP-Kalkulationen eher im Bereich zwischen Aminosäuren 14



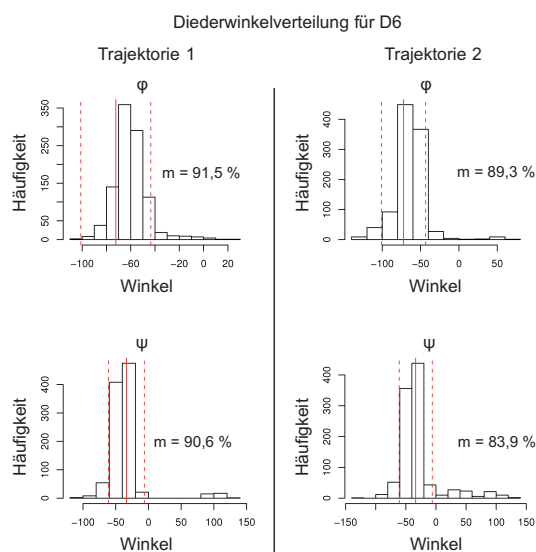
und 18 zu finden ist.

Zusätzlich eingefügt wurden in die Tabelle die aus den Distanzen bestimmten  $\alpha$ -Helices für den Villin-Headpiece 36 Wildtyp 1vii [16]. Verglichen mit  $x_{arithm.}$  weichen auch diese Werte nur ganz geringfügig voneinander ab. Ein signifikanter Unterschied besteht darin, dass die mittlere Helix beim HP36-Wildtyp deutlich definiert ist, was für new2 und mut8 nicht der Fall ist.

Insgesamt lässt sich der Schluss ziehen, dass – bis auf die mittlere Helix – die Sekundärstrukturelemente von new2/mut8 mit denen des HP36-Wildtyps übereinstimmen. Dies bedeutet wiederum, dass die acht Mutationen keine wesentlichen Auswirkungen auf die Sekundärstruktur haben.

### 4.1.3. Die Diederwinkel von mut8 und new2

Nach der Bestimmung der Sekundärstrukturanteile in mut8 über NMR-Daten wurden die Diederwinkel mit Hilfe von TALOS ermittelt (siehe Kap. 3.1.2). Diese Winkelinformationen  $\phi$  und  $\psi$  wurden von Manuel Prinz ebenfalls für die 2000 simulierten Strukturen aus den beiden Trajektorien für new2 erhoben, jeweils gemittelt und mit den entsprechenden  $\phi$ - und  $\psi$ -Werten aus den von TALOS ermittelten Daten für mut8 verglichen.



**Abbildung 4.4.:** Darstellung der Diederwinkelbestimmung von Aspartat 6 in Trajektorie 1 und 2. Die durchgezogene rote Linie beschreibt den von TALOS bestimmten Wert, die gestrichelten roten Linien die erlaubten Grenzen. Der Anteil  $m$  an Strukturen, die sich in der Trajektorie innerhalb des von den beiden gestrichelten Linien eingegrenzten Bereiches befinden, ist in % angegeben.

Abbildung 4.4 zeigt beispielhaft die Bestimmung der Diederwinkel von Aminosäure D6 aus new2 für die Strukturen aus der ersten und zweiten Trajektorie. Eingezeichnet sind die entsprechenden TALOS-basierten Winkel (durchgezogene rote Linie) für



**Abbildung 4.5.:** Darstellung der new2/mut8-Sequenz für Trajektorie 1 und 2. Blau unterlegt sind die Aminosäuren, dessen  $\phi$ - und  $\psi$ -Werte starke Ausreißer darstellen.

mut8 mit den jeweiligen erlaubten Grenzen (rot gestrichelte Linien). m (in %) ist der Anteil an Strukturen, die sich in der Trajektorie innerhalb des nach TALOS gegebenen Toleranzbereiches befinden. Alle übrigen bioinformatischen Auswertungen zu den Diederwinkeln befinden sich im Anhang.

Insgesamt befinden sich so 63 % der Strukturen aus Trajektorie 1 auf den nach TALOS erlaubten  $\phi$ -Winkel-Werten und 50 % dieser Strukturen auf den erlaubten  $\psi$ -Winkel-Werten. Für Trajektorie 2 gilt: 59 % der Strukturen befinden sich im Bereich der nach TALOS erlaubten  $\phi$ -Winkel-Werte und 47 % befinden sich im Bereich der erlaubten  $\psi$ -Winkel-Werte.

Diese teilweise geringen Werte scheinen zunächst für eine mäßig gute Übereinstimmung der bioinformatischen Daten mit denen aus der NMR zu sprechen. Bei näherer Betrachtung  $\phi$ - und  $\psi$ -Werte befinden sich jedoch die Ausreißer, welche das Ergebnis verschlechtern, entweder ganz N- oder ganz C-terminal (Aminosäuren 2 und 3 bzw. 33 und 34) oder mittig bei Aminosäuren 11, 12 bzw. 17 bis 20 (siehe Abb. 4.5).

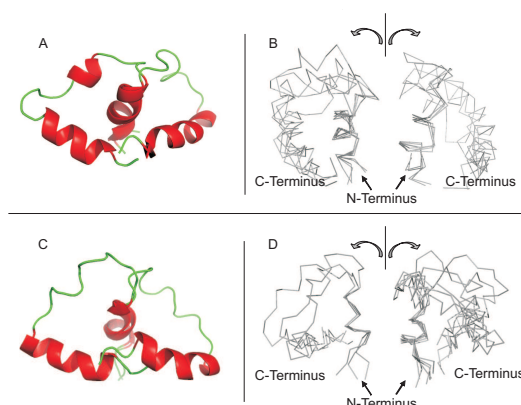
Interessanterweise beschreiben diese letztgenannten Aminosäuren (17 bis 20) den nicht genau definierbaren Mittelteil von mut8 (vgl. Tabelle 4.1), somit ist für diesen Bereich auch eine schlechtere Diederwinkel-Vorhersage über TALOS zu erwarten, woraus sich zwangsweise diese mäßig gute Übereinstimmung mit den bioinformatisch ermittelten Werten ergibt.

Nach der Auswertung der bioinformatisch ermittelten  $\phi$ - und  $\psi$ -Werte wurde ein weiteres Experiment durchgeführt. Dafür wurden die über TALOS ermittelten Diederwinkeln durch die aus der Bioinformatik bestimmten Diederwinkel im *sa.inp*- und *ref.inp*-Protokoll ersetzt, zusammen mit den NOE-Distanzbeschränkungen aus den NMR-Messungen (siehe Kap. 3.1.3) kombiniert und unter Xplor-NIH eine Schar aus 100 neuen Strukturen berechnet (siehe Kap. 2.6.3).

Aus diesen 100 neuen Strukturen wurden die 10 besten (minimale Energie, keine Verletzungen) ausgewählt und mit dem ursprünglichen Ensemble für mut8 (siehe Kap. 3.1.3) verglichen. Durch dieses Experiment kann überprüft werden, ob sich trotz unterschiedlicher Diehedralwinkelsätze dieselben bzw. ähnliche Strukturen aus den Rechnungen ergeben.

Abbildung 4.6 zeigt ein Alignment der 10 besten Strukturen (links) im Vergleich zu dem bereits bekannten Ensemble für mut8 (rechts, vgl. Abb. 3.15).

In Abbildung 4.6 lässt sich deutlich erkennen, dass nach Einsetzen der bioinformatisch berechneten Winkel die Sekundärstrukturelemente in den Strukturen des neues Ensembles fast identisch sind zu denen des vorangehenden Ensembles, sobald die über TALOS bestimmten Werte eingesetzt werden.



**Abbildung 4.6.:** Darstellung der für mut8 berechneten Strukturen des Ensembles im Vergleich zu den Strukturen für r1g11i, welche mit den bioinformatisch ermittelten  $\phi$ - und  $\psi$ -Winkeln berechnet wurden (vgl. auch Abb. 3.15). (A und B) Repräsentative *cartoon*-Darstellung zweier möglicher mut8-Strukturen aus dem Ensemble und Aufteilung des Ensembles in zwei verschiedene Cluster. (C und D) Repräsentative *cartoon*-Darstellung zweier möglicher Strukturen aus dem neuen Ensemble, berechnet mit den bioinformatisch ermittelten  $\phi$ - und  $\psi$ -Winkeln und Aufteilung des Ensembles in zwei Cluster. Alle Strukturen wurden entlang der  $C_{\alpha}$ -Atome der ersten  $\alpha$ -Helix (Aminosäuren 3 bis 11) überlagert.

So findet sich sowohl zu 100 % eine vordere  $\alpha$ -Helix als auch zu 100 % eine hintere, die mittlere  $\alpha$ -Helix ist jedoch bei keiner der Strukturen aus dem neu berechneten Ensemble zu sehen. Da dies aber bei über 50 % der Strukturen, welche mit Winkeln aus der TALOS-Vorhersage berechnet wurden, ebenfalls der Fall ist und die mittlere Helix stets schlechter definiert ist als die beiden anderen, kann dieser Fakt als weniger relevant angesehen werden.

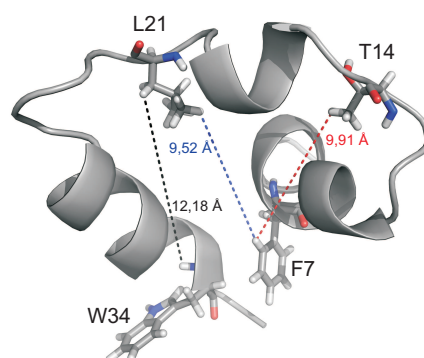
Sowohl die Daten zu den NOEs (NMR/Bioinformatik) als auch die des CSI, der DSSP und der Diederwinkel (NMR/Bioinformatik) miteinander vergleichend, lässt sich sagen, dass sich die Sekundärstrukturverteilung von new2 fast identisch zu der von mut8 verhält, die *in silico*-Auswertung stimmt also mit den aus NMR-Daten gewonnenen Ergebnissen überein.

## 4.2. Die Tertiärstruktur von mut8

Villin Headpiece 36 (HP36) dient als Modell in der Bioinformatik, um Molekuldynamik-Simulationen und Faltungsstudien durchzuführen (siehe Kap. 1). HP36 stellt mit seinen 36 Aminosäuren ein kleines, unabhängig faltendes Peptid mit sehr stabiler Tertiärstruktur dar und ist seit den späten 1970er-1990er Jahren sowohl in seiner Struktur als auch in seiner Funktion gut untersucht [10, 12, 16].

Neben dem Wildtyp des HP36 wurden auch Mutanten und Verwandte untersucht. Aus den Ergebnissen dieser Untersuchungen wurde ersichtlich, dass sich alle Mutan-





**Abbildung 4.9.:** Darstellung einer mut8-Struktur des Ensembles im *cartoon*-Modell mit Hilfe von PyMol. Hervorgehoben sind Aminosäuren F7, T14, L21 und W34. Zusätzlich eingezeichnet sind die nach PyMol bestimmten Abstände zwischen  $7H_{\delta^*}/14H_{\gamma 1}$  (rot),  $7H_{\delta^*}/21H_{\delta^*}$  (blau) und  $21H_{\beta^*}/34H_N$  (schwarz).

ne Übereinstimmung der Tertiärstruktur der Achtfachmutante zu 1vii erwarten. Dies war jedoch nicht der Fall; so führte der Austausch der acht Aminosäuren zwar nicht zur vollständigen Destabilisierung, da Sekundärstrukturelemente wie  $\alpha$ -Helices immer noch nachweislich vorhanden waren und die aufgenommenen Spektren für eine Strukturierung sprachen, jedoch mussten mut8 nun dynamische Eigenschaften zugeschrieben werden, die der Wildtyp vorher nicht aufwies.

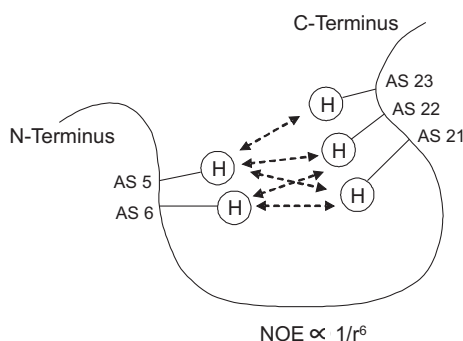
Die Tertiärstruktur eines über NMR-Spektroskopie vermessenen Proteins oder Peptids wird maßgeblich durch die Anzahl der weit reichenden NOEs bestimmt. Diese und alle anderen NOEs (intraresiduale, sequentielle, mittelweit reichende) können sowohl manuell aus dem NOESY-Spektrum ermittelt als auch mit Hilfe der Algorithmen aus dem Programm AUREMOL [78] automatisch zugeordnet werden. Eine Kombination aus beiden Möglichkeiten ergibt dabei die effizienteste Strategie zur Zuordnung aller Signale.

Die zugeordneten NOEs und die über TALOS ermittelten Diederwinkel wurden als experimentelle Parameter in Xplor-NIH zur Tertiärstrukturberechnung eingesetzt (siehe Kap. 3.1.3).

Die Bestimmung einer Tertiärstruktur für mut8 erwies sich jedoch als wesentlich schwieriger als die Auswertung der stabilen Sekundärstrukturelemente. Nicht alle NOE-Distanzbeschränkungen konnten gleichzeitig erfüllt werden, d. h. es waren NOEs im NOESY zu sehen, welche sich aus energetischen Gründen in einer gefalteten Tertiärstruktur von mut8 gegenseitig ausschlossen.

Abbildung 4.9 zeigt beispielhaft eine der kalkulierten mut8-Strukturen aus dem Ensemble als *cartoon*-Modell. Hervorgehoben sind die Seitenketten des Phenylalanin 7, Threonin 14, Leucin 21 und Tryptophan 34. Mit eingezeichnet sind die für diese Struktur berechneten Abstände zwischen  $7H_{\delta^*}/14H_{\gamma 1}$  (rot),  $7H_{\delta^*}/21H_{\delta^*}$  (blau) und  $21H_{\beta^*}/34H_N$  (schwarz).

Im NOESY-Spektrum sind für diese Abstände Signale zu sehen, für welche aufgrund



**Abbildung 4.10.:** Schematische Darstellung von dipolaren Kopplungen zwischen Protonen der Aminosäuren 5 und 6 (AS 5, 6) und Aminosäuren 21 bis 23 (AS 21, 22, 23) innerhalb einer Peptidkette. Der aus den Kopplungen resultierende NOE ist abhängig vom Abstand  $r$  zwischen den beiden Protonen ( $1/r^6$ ). Verändert nach [113].

der NOE-Intensitäten Abstände zwischen 3,5 und 4,8 Å vorschlagen würden. Nach Eingabe aller zu sehenden, korrekt und eindeutig zugeordneten NOEs war es nicht möglich, mit Xplor-NIH realistische Tertiärstrukturen zu berechnen, welche Energien unter 100 kcal/mol besitzen und gleichzeitig keine Abweichungen von den experimentellen Daten (keine NOE-Verletzungen, keine Verletzungen der gesetzten Diederwinkel) aufweisen.

Um dennoch ein Ensemble aus Strukturen berechnen zu können, wurden die NOEs selektiert. Die Wahrscheinlichkeit, dass weitreichende NOEs korrekt sind, wird dadurch erhöht, dass diese nicht vereinzelt auftreten. Das bedeutet, dass mehrere NOEs von Protonen einer Aminosäure  $i$  zu Protonen einer Aminosäure und ihren direkten Nachbarn  $i+x$ ,  $i+(x+1)\dots i+(x+n)$  für  $x \geq 4$  erwartet würden (siehe Abb. 4.10).

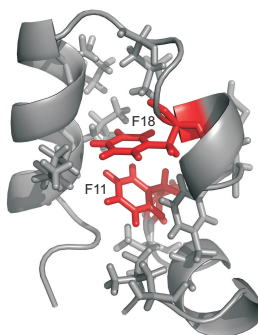
Da NOEs zwischen Aminosäuren 7 und K30/Y33/W34/L35/F36 zu sehen waren, wurde von der Richtigkeit dieser Zuordnung ausgegangen. Xplor-NIH gibt nach jeder Strukturberechnung an, welche der gesetzten NOEs verletzt werden. So wurden nach und nach alle verletzten NOEs entfernt, welche mit den anfangs eingesetzten NOEs zwischen Aminosäuren F7 und K30, Y33, W34, L35, F36 nicht kompatibel waren. Weitere akzeptable NOEs waren die zwischen L21 und Q26/L29. Obwohl die meisten dieser NOEs nur sehr schwach im Spektrum zu sehen waren, konnte durch diese ein Ensemble an Strukturen berechnet werden (siehe Kap. 3.1.3). Dennoch müssen die Strukturen des Ensembles mit Vorsicht und nur als Möglichkeit einer Tertiärstruktur für mut8 betrachtet werden, da, wie bereits erwähnt, einige genau definierbare NOE-Distanzbeschränkungen existieren, jedoch nicht zur Strukturberechnung eingesetzt wurden.

Die Literatur beschreibt ebenfalls zahlreiche Villin Headpiece 36-Mutanten (siehe Kap. 4), welche jedoch alle stabil falten und somit eine geordnete Tertiärstruktur ergeben. Dennoch handelt es sich dabei meist um Einfach- bzw. Zwei- oder Dreifachmutanten, wohingegen bei mut8 acht Aminosäuren gegenüber dem Wildtyp ausgetauscht wurden. Dabei kommt es auch darauf an, welche Aminosäuren mutiert werden. Ungefähr 10 Aminosäuren sind an der Bildung eines hydrophoben Kerns in Villin Headpiece



36 (1vii) beteiligt [24] (siehe auch Kap. 1): L42 (Nummerierung nach 1vii), F7, V10, F11, M13, F18, L21, K25, L29 und K30 (Nummerierung nach mut8/new2). Natürliche Mutationen treten selten an Positionen 7, 11, 18 und 21 auf, was darauf hindeutet, dass diese Aminosäuren für die charakteristische Faltung von HP36 wichtig sind.

Abbildung 4.11 zeigt die Darstellung des Villin Headpiece 36-Wildtyps (1vii) mit den den hydrophoben Kern ausmachenden Aminosäuren als *stick*-Darstellung hervorgehoben, die in mut8 Mutierten sind in rot dargestellt.



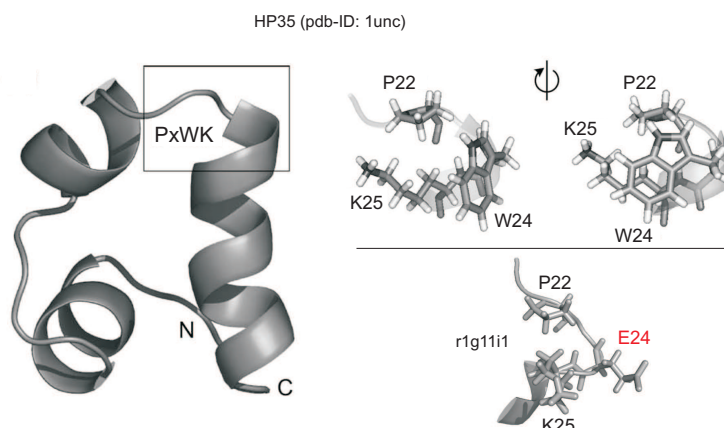
**Abbildung 4.11.:** Darstellung von 1vii. Hervorgehoben sind alle für den hydrophoben Kern wichtigen Aminosäuren sowie die beiden in r1g11i1 mutierten Residuen F11D und F18K in rot.

Eine höhere Toleranz gegenüber Mutationen an Positionen 10, 13 und 29 deutet darauf hin, dass diese Aminosäuren lediglich zusätzlich zur Stabilisierung der Faltung beitragen. Vermeulen *et al.*, 2004 beschreiben weiterhin, dass das Glutamin an Position 26 (Nummerierung nach mut8/new2) des HP36-Wildtyps „eingeklemt“ zwischen mittlerer und letzter Helix liegt und das  $H_{\epsilon 22}$  dieser Aminosäure eine Wasserstoffbrücke mit dem Rückgrat-CO von L21 bildet. Tatsächlich konnten im NOESY-Spektrum von mut8 auch drei weitreichende NOEs zwischen Aminosäuren 21 und 26 ermittelt werden ( $21H_{\delta 1^*}/26H_{\alpha}$ ,  $21H_{\beta 2}/26H_{\epsilon 2^*}$  und  $21H_{\beta^*}/26H_N$ ), die darauf hinweisen, dass sich Aminosäuren 21 und 26 in räumlicher Nähe zueinander befinden.

Sowohl F11 als auch F18, für dessen Austausch destabilisierende Effekte beschrieben sind [114], sind in mut8 mutiert (F11D, F18K). Gleichzeitig erwähnen Xiao *et al.*, 2009, dass zwar ein Austausch jeglicher Phenylalanine in 1vii zu einer Abnahme der strukturellen Stabilität führt, jedoch selbst ein Austausch aller drei in der Sequenz von 1vii vorkommenden Phenylalanine durch Leucine nicht dazu führt, dass das Peptid sich nicht mehr korrekt faltet.

Das einzige in HP36 vorkommene Prolin (P22) spielt für die Faltung des Peptids ebenfalls eine entscheidende Rolle. So kann dieses sich nicht mehr nativ falten, sobald P22 gegen ein Alanin ausgetauscht wurde [25, 23].

Ebenfalls wurde gezeigt, dass das konservierte Tryptophan an Position 24 eine kritische Rolle bei der Faltung von HP36 spielt, da es mit dem P22 interagiert [25]. Für solche W-P-Interaktionen ist allgemein bekannt, dass sie sich stabilisierend auf die Struktur von kleinen Domänen auswirken [115, 116, 117, 118]. Hydrophobe Berei-



**Abbildung 4.12.:** Darstellung von Villin Headpiece 35 (pdb-ID: 1unc) mit dem PxWK-Motiv (links). Die Positionen der Seitenketten von P22, W24 und K25 in PxWK sind rechts oben gezeigt, rechts unten der entsprechende Teil aus mut8, in dem W24 gegen ein Glutamat ausgetauscht ist. Verändert nach [25].

che auf der Oberfläche von Proteinen können dazu beitragen, dessen gefaltete Struktur zu stabilisieren [119]. Diese Tatsache führte dazu, den evolutionär konservierten, Lösungsmittel-exponierten hydrophoben Bereich von HP36, das PxWK-Motiv (Aminosäuren 22 bis 25), näher zu untersuchen. In diesem Motiv liegt der aromatische Indolring des W24 über den aliphatischen Seitenketten der hochkonservierten P22 und K25 (siehe Abb. 4.12), eine räumliche Anordnung, die in alle bisher gelösten Villin Headpiece-Strukturen zu finden ist [25]. In mut8 ist W24 gegen ein Glutamat ausgetauscht, was zur Folge haben könnte, dass das Peptid zusätzlich destabilisiert wird.

Xiao *et al.*, 2009 beschreiben, dass für die Kopplung zwischen Aminosäuren L2 und L35 sowie F11 mit K33 NOEs existieren. Diese Interaktionen sind bei mut8 nicht mehr vorhanden, L2 wurde gegen ein Tyrosin ausgetauscht sowie das hydrophobe F11 gegen ein polares Aspartat und das stark hydrophobe K33 gegen ein Tyrosin.

Auffällig ist weiterhin, dass bei der Tertiärstrukturberechnung des HP36-Wildtyps (1vii) 69 weitreichende NOEs eingesetzt wurden [17] und bei der Strukturaufklärung des humanen HP36 (1unc) sogar 145 weitreichende NOEs gefunden wurden [24]. Im Gegensatz dazu wurden zur Strukturberechnung von mut8 lediglich 19 weitreichende NOEs gefunden. Dies ist ein weiterer Hinweis darauf, dass die Tertiärstruktur von mut8 nicht eindeutig bestimmbar ist.

Der Carbonyl-Sauerstoff der Aminosäure F18 steht über eine Wasserstoffbrückenbindung mit dem Rückgrat-NH von L21 in Kontakt. Diese H-Brücke definiert die Schleife zwischen mittlerer und letzter Helix in HP36, welche wiederum wichtig bei der initialen Faltung des Peptids ist [120, 20]. In mut8 ist F18 gegen ein Lysin ausgetauscht. Auch wenn die H-Brücke zwischen Rückgratatomen erfolgt, so könnte dennoch ein solcher Austausch dazu führen, dass diese Bindung nicht mehr vorhanden ist und somit die initiale Faltung in mut8 fehlt.



Die Lösungsmittel-exponierte Position 12 ist mit einem Glycin belegt, der flexibelsten Aminosäure. Das flexible G12 kann dazu beitragen, dass die erste Helix an die mittlere und letzte Helix docken kann. Eine Mutation an dieser Stelle könnte somit die konformationelle Freiheit der dort in der Struktur von HP36 befindlichen Schleife reduzieren [121]. In diesem Fall ist das Glycin an Stelle 12 tatsächlich erhalten geblieben.

Der Algorithmus EMOS führt einen *in vitro alanine scan* durch (vgl. Kap. 1.2). Bei diesem wird überprüft, ob Aminosäuren gegen ein Alanin austauschbar sind, ohne dass es z. B. zu einer Destabilisierung des Peptids kommt. Interessanterweise hat EMOS G12 und P22 als wichtige Aminosäuren erkannt und nicht mutiert, während Aminosäuren F11, W24, K33 ausgetauscht und die initiale F18K-Mutation (siehe Kap. 1) beibehalten wurden, obwohl diese als wahrscheinlich wichtig für die Faltung von mut8 zu deklarieren sind. Dies könnte ein wichtiger Anhaltspunkt dafür sein, dass der Algorithmus noch nicht vollständig ausgreift ist.

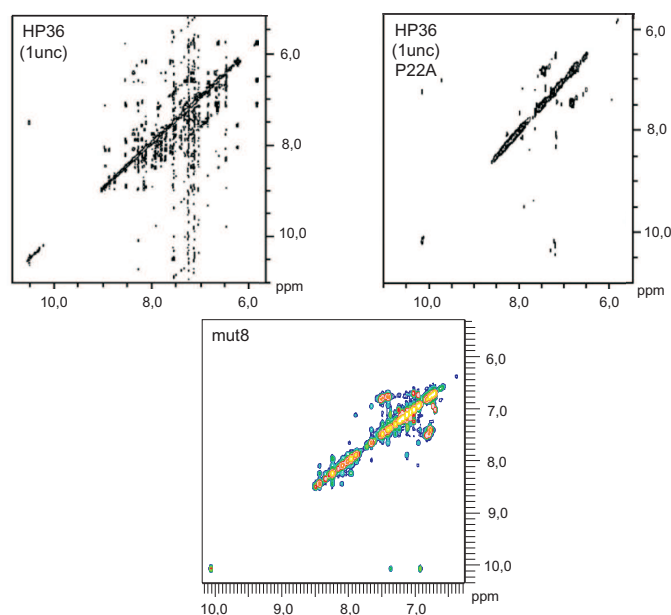
Aufgrund dieser Tatsachen kann nun darauf geschlossen werden, dass die Kombination aus Destabilisierung der Struktur durch Austausch von F11 und F18 und Austausch weiterer sechs Aminosäuren (L2Y, D4A, W24E, E32A, K33Y und G34W) zu einer Auflösung oder extrem starken Destabilisierung der Tertiärstruktur von mut8 führt. Dies könnte eine mögliche Erklärung für die Dynamik dieses Peptids sein.

### 4.3. Expression von mut8 und heteronukleare NMR-Spektroskopie

Während der NMR-Analysen wurde deutlich, dass zwar die Sekundärstrukturelemente in mut8 definiert werden konnten, sich einige NOEs bei der Berechnung einer Tertiärstruktur jedoch nicht gleichzeitig miteinander kombinieren ließen (vgl. Abb. 4.9). Dies wies darauf hin, dass nicht nur eine Konformation von mut8 vorliegt (siehe Kap. 3.1.1), sondern mehrere Strukturen nebeneinander existieren.

Die Tatsache, dass die chemisch synthetisierte Form von mut8 in wässriger Lösung nicht aggregiert und auswertbare Sekundärstrukturelemente aufweist, ließ vermuten, dass es sich um ein dynamisches Peptid handeln könnte. Dabei gehen die verschiedenen Konformationen so schnell ineinander über und existieren gleichzeitig in Lösung, dass in einem NOESY die NOEs aus den gemittelten Konformationen zu sehen sind. Da die einzelnen Konformere auf der Zeitskala der chemischen Verschiebung nicht sichtbar waren (keine doppelten Signale im HSQC und keine Linienverbreiterungen erkennbar), muss davon ausgegangen werden, dass sich die Dynamik im ns- bis  $\mu$ s-Bereich abspielt.

Bei Peptiden, welche ein  $\omega\tau_c \approx 1$  besitzen, ist der maximale homonukleare NOE ( $\eta$ )=0. Eine kleine Rotationskorrelationszeit  $\tau_c$  ist bei kleinen Peptiden gegeben, wodurch  $\omega\tau_c$  gegen 1 gehen kann (siehe Abb. 3.23). Möglicherweise ist dies bei mut8 der Fall, sodass  $\eta=0$  wird und dadurch nicht alle NOEs im Spektrum sichtbar sind. Dass wesentlich weniger NOEs im Spektrum zu sehen sind als beim HP36-Wildtyp, zeigt Abbildung 4.13. Zwar sind dort im Amidbereich Signale zu erkennen, jedoch wesentlich weniger als beim Wildtyp. Eine geringe Signaldichte im Amidbereich zeigt auch



**Abbildung 4.13.:** Die Amidregion der NOESY-Spektren von HP36 (1unc) (links), von dessen HP36-P22A-Mutante (rechts) und von mut8. Deutlich zu erkennen ist die Verringerung der Anzahl der Signale vom Wildtyp zu den Mutanten. Verändert nach [23].

das NOESY-Spektrum für die P22A-Mutante des HP36, einer Mutante, die nachweislich nicht mehr korrekt faltet [25, 23]. Diese Parallele ist ebenfalls ein Hinweis darauf, dass es bei mut8 nicht mehr zur Ausbildung einer stabilen Tertiärstruktur kommt.

Die Klonierung erfolgte in einen pET-Vektor mit GST-*tag* (siehe Kap. 3.2). Exprimiert wurde das GST\_mut8-Konstrukt in *E. coli* Rosetta (DE3) in  $^{15}\text{N}$ -markiertem M9-Medium. Nach der Expression und Reinigung von mut8 wurde deutlich, dass das Konstrukt noch während der Expression einem proteolytischen Abbau unterlag (vgl. Kap. 3.3.1), der trotz Zugabe von Proteasehemmern nicht zu minimieren war. Im SDS-Gel waren neben der Hauptbande mehrere mut8-Abbaubanden sichtbar (siehe Abb. 3.20, links).

Dass der Abbau am C-Terminus stattfand, konnte dadurch festgestellt werden, dass im  $^{15}\text{N}$ -HSQC kein charakteristisches  $\text{H}_{\epsilon 1}$ -Signal des W34 bei 10,06 zu sehen war. Der N-Terminus war durch den GST-*tag* vor Abbau geschützt. Um den Abbau zu stoppen bzw. zu minimieren, kamen zwei Möglichkeiten in Frage: Entweder konnte C-terminal auch ein *tag* angebracht werden oder es musste die Expressionsmethode selbst verändert werden. Die erste Möglichkeit wurde ausgeschlossen, da alle verfügbaren *tags* nach Abschneiden über z. B. der PreScission-Protease (Erkennungssequenz FFQ|GP) mindestens zusätzliche 3 Aminosäuren an den C-Terminus addieren würden, obwohl bereits am N-Terminus zusätzlich zur mut8-Sequenz ein Glycin und ein Prolin zu finden sind. Diese drei nicht-nativen Aminosäuren würden zusammen mit dem bestehenden N-terminalen GP das Peptid um 14 % verlängern und könnten die Sekundär-

und Tertiärstruktur des relativ kleinen mut8 stark verfälschen oder zumindest Einfluss darauf nehmen.

Daher wurde beschlossen, die Expressionsmethode zu verändern. Die Zellen wurden bis zu einer OD<sub>600</sub> von 1,5 angezogen und dann nur 5 h lang bei 28 °C exprimiert (siehe auch Kap. 2.3.3). Dadurch konnte die Anzahl an exprimierenden *E. coli*-Zellen erhöht, die Angriffszeit und möglicherweise die Menge an endogenen Proteasen aber vermindert werden. Tatsächlich konnte durch diese Kurzexpressionsmethode der Abbau zwar nicht gestoppt, jedoch sichtlich verringert werden (vgl. Abb. 3.20, rechts).

Im entsprechenden <sup>15</sup>N-HSQC des so gereinigten Peptids war das H<sub>ε1</sub>-Signal des W34 nun in doppelter Form zu sehen, obwohl nur ein Tryptophan in der Sequenz vorhanden ist. Da jedoch die Signale für L35 und F36 eindeutig zugeordnet werden konnten, musste davon ausgegangen werden, dass neben dem vollständigen mut8 auch mindestens ein C-terminales Abbauprodukt vorhanden war. Dass sich mut8 in genau zwei verschiedenen, nur langsam ineinander übergehenden Konformationen zugleich aufhält und daher zwei H<sub>ε1</sub>-Signale des W34 zu sehen waren, war auch hier unwahrscheinlich. Dies konnte dadurch ausgeschlossen werden, dass nach 11 Tagen Lagerung der Probe bei 8 °C weitere schwächere Signale auftraten und ein drittes H<sub>ε1</sub>-Signal des W34 im Spektrum sichtbar wurde. Um ausschließen zu können, dass es sich um zwei langsam ineinander übergehende Konformationen handelt, hätte eine Temperaturreihe zu geringeren und höheren Temperaturen durchgeführt werden können. Bei zwei Konformationen müssten die entsprechenden zwei auftretenden Signale der jeweiligen Aminosäure dann im Spektrum aufeinander zulaufen und schließlich zusammenfließen. Aufgrund des weiteren zeitlichen Abbaus von mut8 wurde jedoch darauf verzichtet.

Da der C-terminale Abbau nicht gänzlich gestoppt werden konnte, wurde mit der durch Abbauprodukte verunreinigten Probe weiter gearbeitet. Um die Signale für das vollständige Peptid von den Signalen der Abbauprodukte zu unterscheiden, wurde ein HMQC des chemisch synthetisierten mut8 zu Referenzzwecken aufgenommen. Da diese Probe jedoch nicht <sup>15</sup>N-markiert war und die Datenakquisition daher mit einem herkömmlichen HMQC-Pulsprogramm über eine Woche gedauert hätte, wurde ein SOFAST-HMQC (vgl. 2.5.3) innerhalb von 21 h aufgenommen (vgl. Abb. 3.22). Die Pulssequenz dieses Experimentes ist so optimiert, dass nur sehr kurze Verzögerungszeiten  $t_{rec}$  zwischen den einzelnen *scans* entstehen und somit die Datenakquisition wesentlich schneller erfolgen konnte [86, 122].

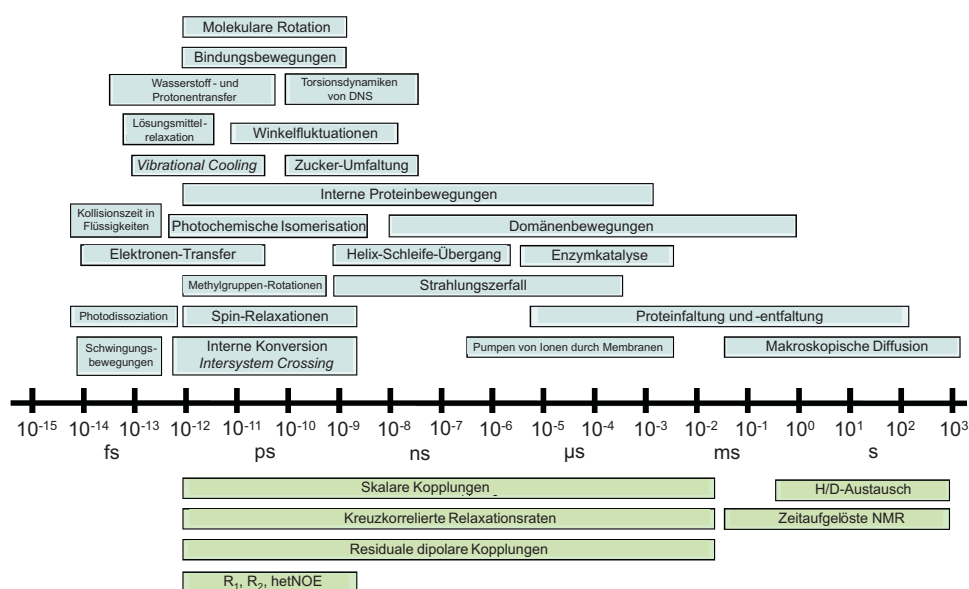
Die Zuordnung der SOFAST-HMQC-Signale erfolgte daraufhin mit Hilfe der homonuklearen Spektren des chemisch synthetisierten mut8. Es konnten die H<sub>N</sub>-Signale aller Aminosäuren mit Ausnahme von M1 und Y2 sowie P22 aufgrund des nicht zugeordneten bzw. fehlenden H<sub>N</sub> im Spektrum und den Aminosäuren M13, A17, K18, K25, L29 und K30 zugeordnet werden. Letztere zeigten im Spektrum zu dicht aneinanderliegende oder überlappende Amidprotonenfrequenzen (vgl. Abb. 3.22).

## 4.4. Bei mut8 handelt es sich um ein dynamisches Peptid

Nach der Aufnahme des SOFAST-HMQC konnte mit Hilfe der rekombinant exprimierten Probe die Dynamik des Peptids untersucht werden. Dazu wurden HSQCNOE-Spektren aufgenommen.

### hetNOE und RMSF

HSQCNOE-Spektren geben Auskunft über die Bewegungen von  $H_N$ -Vektoren im ps- bis ns-Bereich (vgl. Abb. 4.14).



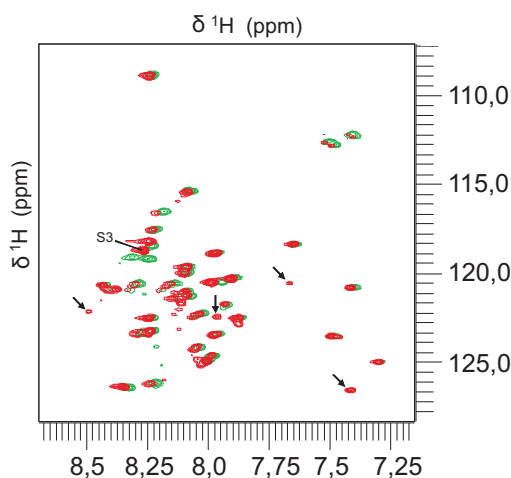
**Abbildung 4.14.:** Darstellung der Zeitskala für molekulare Dynamik (blau) und NMR-spektroskopische Methoden (grün) (verändert nach [4]).

Beim HSQCNOE werden zwei Spektren simultan aufgenommen, eines mit Protonensaturierung, eines ohne Protonensaturierung. Die aus diesen Spektren resultierenden Signale besitzen somit die Intensität  $I_{sat}$  bzw.  $I_{eq}$ . Der hetNOE ist der Quotient aus diesen Intensitäten (bzw. Volumina) der Signale beider HSQCNOE:  $I_{sat}/I_{eq}$ .

Die Zuordnung der Signale des SOFAST-HMQC konnte auf das HSQCNOE<sub>eq</sub> übertragen werden (siehe Kap. 3.4.2), sodass tatsächlich nur die Signale zugeordnet werden konnten, welche zu dem vollständigen mut8 gehörten und nicht zu dessen Abbauprodukten. Abbildung 4.15 zeigt die Überlagerung des SOFAST-HMQC (grün) und des HSQCNOE<sub>eq</sub> (rot) (vgl. Kap. 3.4.2).

Mit Pfeilen markiert sind die Resonanzen, welche zusätzlich im SOFAST-HMQC auftauchen und daher wahrscheinlich von den Abbauprodukten des rekombinant ex-

primierten mut8 stammen. Alle Signale des HSQCNOE<sub>eq</sub> weichen um etwa 0,03 bis 0,08 ppm von den Signalen des SOFAST-HMQC ab, was u. a. daran liegen könnte, dass bei der <sup>15</sup>N-markierten Form des mut8 aufgrund des nach der Reinigung abgeschnittenen GST-tags noch ein Glycin und Prolin am N-Terminus vorhanden sind. Je näher sich nun eine Aminosäure der mut8-Sequenz an den zusätzlichen Aminosäuren Glycin und Prolin befinden, desto stärker wird deren chemische Umgebung beeinflusst und desto stärker würde auch das Signal von dem „Ursprungssignal“ (aus dem SOFAST-HMQC) abweichen. Dies ist auch tatsächlich der Fall, so weicht das Signal der Aminosäure S3 mit 0,08 ppm am stärksten von dem äquivalenten Signal aus dem SOFAST-HMQC ab. Da Aminosäuren M1 und Y2 nicht zugeordnet werden konnten, kann für deren möglicherweise im Spektrum sichtbaren Signale keine Aussage getroffen werden.



**Abbildung 4.15.:** Ausschnitt des SOFAST-HMQC (grün, vgl. Abb. 3.22) und des HSQCNOE<sub>eq</sub> (rot). Mit Pfeilen markiert sind die Signale, welche nicht zuzuordnen sind und wahrscheinlich von Abbauprodukten stammen. Markiert ist weiterhin Aminosäure S3, deren Signal mit 0,08 ppm am stärksten zu dem Ursprungssignal im SOFAST-HMQC abweicht.

Die Signale der Aminosäuren M13, A17, K18, K25, L29 und K30, welche nicht genau zugeordnet werden konnten, da deren  $H_N$ -Werte zu nah beieinander lagen, wurden im weiteren Verlauf der Rechnung nicht mit berücksichtigt. Im HSQCNOE<sub>sat</sub>-Spektrum waren die Signale der Aminosäuren T14, R15, S16, N20, Q27, N28 und A32 nicht mehr zu sehen, d. h. dafür wurde bei der Bildung des Quotienten aus  $I_{sat}/I_{eq}$  jeweils der Wert 0 für  $I_{sat}$  eingesetzt, womit sich zwangsweise auch ein Quotient von 0 ergibt. Das Ergebnis der Dynamikberechnung (siehe Abb. 3.24) zeigte, dass vor allem die Bereiche von Aminosäuren 3 bis 6 und 33 bis 36 sehr flexibel sind (negativer hetNOE von -0,06 bis -0,4) sowie ein Mittelteil zwischen Aminosäuren 14 und 20 existiert, welcher mit einem hetNOE=0 mäßig flexibel zu sein scheint.

Für strukturierte Proteine liegt der hetNOE bei Werten zwischen 0,8 und 0,9, klei-

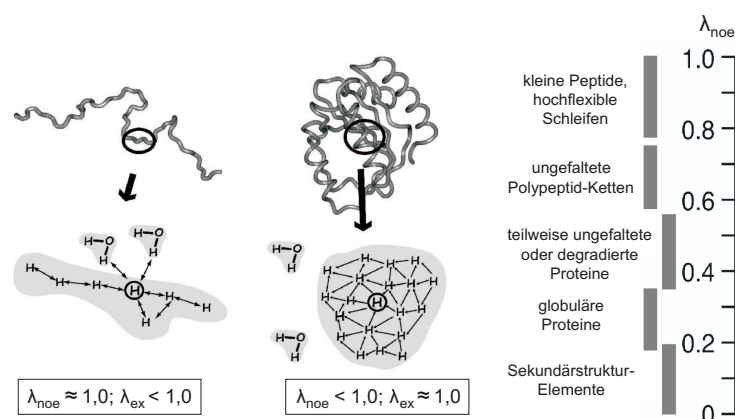
nere Werte sprechen für flexible Bereiche, negative Werte für hochflexible Regionen, wie sie an N- und C-Termini von Proteinen zu finden sind. Der hier vorhandene höchste positive hetNOE liegt bei 0,25, was sich möglicherweise darauf zurückführen lässt, dass mut8 mit 36 Aminosäuren ein sehr kleines Peptid darstellt und daher die interne Dynamik generell größer ist als die von großen Proteinen. Neuhaus und Williamson (2000) beschreiben zudem, dass der hetNOE auch als  $(I_{sat}/I_{eq})-1$  dargestellt werden kann (siehe auch Gleichung 3.3, [105]), somit würden sich für mut8 Werte von -1,4 bis maximal -0,75 ergeben. Die relativ geringen positiven und vorwiegend negativen hetNOE-Werte sind ein Hinweis darauf, dass die Tertiärstruktur instabiler ist und größere dynamische Prozesse im Peptid zu beobachten sind.

In der Literatur wird die Dynamik von größeren, gut strukturierten Peptiden (ab  $\approx 70$  Aminosäuren) und Proteinen erwähnt. So beschreiben beispielsweise Brown *et al.*, 2009 eine durchgehend geringe Dynamik ( $\eta_{max} \approx 0,65$ ) für Supervillin-Headpiece, einer 76 Aminosäuren großen Domäne aus der Villin-Typ Headpiece-Familie [123]. Nur im N-terminalen Bereich (Aminosäuren 1 bis 15) sinkt der hetNOE auf -1,6 ab. Ebenso verhält es sich für Dematin Headpiece 76, für das ein maximaler hetNOE von etwa 0,75 und ein minimaler hetNOE von -0,2 (N-terminal) beschrieben wird [124]. Etwas größere Peptide aus 100 bis 134 Aminosäuren wie dem designten Peptid URN (pdb-ID: 2A3J) (100 AS), einer SH2-Domäne (105 AS) und Villin 14T (134 AS) weisen einen  $\eta_{max}$  von 0,7 bis 0,75 und für N-terminale Bereiche einen  $\eta_{min}$  von -0,7 bis -1 auf [112, 87, 125, 126].

Diese Werte für stabil faltende Peptide stehen im Kontrast zu den für mut8 ermittelten hetNOE-Ergebnissen. Wirmer *et al.*, 2006 beschreiben ähnliche hetNOE-Werte für nicht gefaltete Proteine. So wurde in einem Experiment Ubiquitin (76 AS) mit 8 M Harnstoff bei pH 2 vollständig denaturiert und der hetNOE ermittelt [127]. Es ergaben sich maximale hetNOEs von 0,2 ( $\eta_{min} \approx -2$ ). Dieselben Resultate ergaben sich auch für andere denaturierte Proteine [128, 129], so verhält sich die destabilisierte SH3-Domäne mit einem  $\eta_{max}=0,2$  und das Pro-Peptid von Subtilisin (77 AS) mit einem  $\eta_{max}=0,1$  wie das denaturierte Ubiquitin und somit auch wie die instabile HP36-Mutante mut8 ( $\eta_{max}=0,24$ ).

Relativ viele eukariotische Proteine (30-60 %) liegen überwiegend unstrukturiert vor [130, 131]. Ergebnisse aus Proteom-Projekten, die sich mit der Aufklärung der Struktur von Proteinen beschäftigen, zeigen, dass eine erhebliche Anzahl an exprimierten Proteinen wahrscheinlich teilweise ungefaltet vorliegt. Interessante Methoden zur Bestimmung der Dynamik von sowohl gefalteten als auch ungefalteten Proteinen beschreiben Schanda *et al.*, 2006 [132]. Sie verwendeten die von ihnen neu entwickelten 1D- $^1\text{H}$ -HETerogeneity-SOFAST und 2D- $^1\text{H}$ - $^{15}\text{N}$ -HETerogeneity-SOFAST-HMQC, um Informationen über die Gesamtdynamik des untersuchten Proteins zu erhalten (1D-Experiment) oder um Informationen über lokale dynamische Bewegungen zu ermitteln (2D-Experiment).

Im 1D-Experiment erhält man ein globales  $\lambda_{noe}$ , das eine Aussage über die Protonendichte liefert und ein globales  $\lambda_{ex}$ , über welches eine Aussage über die allgemeine Lösungsmittel-Zugänglichkeit der Amidprotonen getroffen werden kann. Damit erhält man einen vollständigen Datensatz zur strukturellen Kompaktheit des untersuchten Proteins. Im 2D-Experiment werden die  $\lambda_{ex}$ - und  $\lambda_{noe}$ -Werte für jedes Signal ermittelt,



**Abbildung 4.16.:** Darstellung der  $^1\text{H}$ - $^1\text{H}$  Spin-Diffusion und von Amid-Wasser-Protonen Austauschwirkungen in einer ausgestreckten Peptidkette und einem globulären, strukturierten Protein. Die  $\lambda_{noe}$ -Werte geben Auskunft über das lokale  $^1\text{H}$ -Spin-Netzwerk, also die Anzahl, Dichte und Beweglichkeit von Protonen, die sich in der Nähe eines Amids befinden. Die  $\lambda_{ex}$ -Werte zeigen, wie Wasser-zugänglich die Amidprotonen sind (links). Charakteristische  $\lambda_{noe}$ -Werte für Proteine unterschiedlicher Struktur und Kompaktheit (rechts). Verändert nach [132].

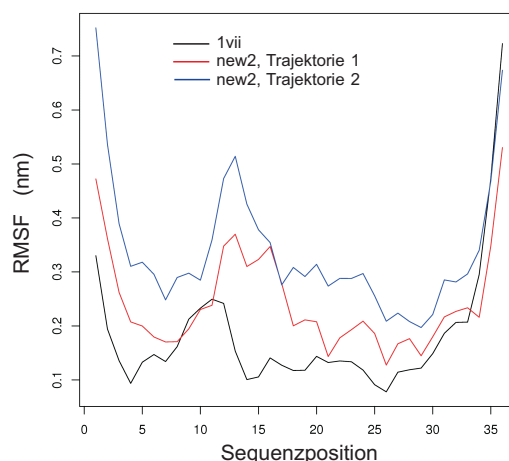
woraus das arithmetische Mittel berechnet wird für die Bestimmung der strukturellen und dynamischen Eigenschaften entlang der Peptidkette.

Abbildung 4.16 (links) zeigt, wie in einem globulären Protein ein Proton von vielen anderen Protonen umgeben ist, jedoch nicht leicht von Wassermolekülen erreicht werden kann. Dadurch erhält man gute  $^1\text{H}$ -Spin-Diffusion ( $\lambda_{noe} \ll 1$ ), aber nur wenig Protonenaustausch ( $\lambda_{ex} \approx 1$ ). Das Gegenteil gilt dafür, wenn eine ausgestreckte Peptidkette untersucht wird. Dort existiert eine niedrige Protonendichte ( $\lambda_{noe} \approx 1$ ) und eine hohe Lösungsmittel-Zugänglichkeit ( $\lambda_{ex} \ll 1$ ). In Abbildung 4.16 (rechts) sind mit Hilfe des 1D- $^1\text{H}$ -HET-SOFAST ermittelte charakteristische  $\lambda_{noe}$ -Werte für verschiedene Proteine unterschiedlicher Struktur, Größe und Kompaktheit dargestellt.

Der Vorteil an dem 1D- $^1\text{H}$ -HET-SOFAST-Experiment ist, dass man in nur 10 s bei einer Proteinkonzentration von 2 mM Spektren aufnehmen kann und gleichzeitig kein  $^{15}\text{N}$ -markiertes Protein mehr benötigt. Ist letzteres dennoch vorhanden, so können über die 2D-Variante zusätzliche, detailliertere Informationen zur Dynamik eines Proteins ermittelt werden. Diese Kombination aus Experimenten bietet sich hervorragend an, um Dynamikstudien von ungefalteten Peptiden wie mut8 durchzuführen.

Um das Ergebnis über die Dynamik des Peptids zu verifizieren, wurde von Manuel Prinz sowohl eine RMSF (*Root Mean Square Fluctuation*)-Auswertung der beschriebenen Trajektorien für new2 als auch für den Wildtyp (1vii) als Referenz durchgeführt.

Über die RMSF erhält man Informationen zur lokalen strukturellen Flexibilität, thermalen Stabilität und Heterogenität von Makromolekülen [133, 134, 135]. Die RMSF ist die Standardabweichung einer Atom- zu einer Referenzposition in nm, gemittelt über die Zeit  $t$  der jeweiligen Trajektorie. Für diese Arbeit wurde die RMSF



**Abbildung 4.17.:** Darstellung der RMSF-Werte für den HP36-Wildtyp 1vii (schwarz) und new2 aus 2 Trajektorien (rot, blau).

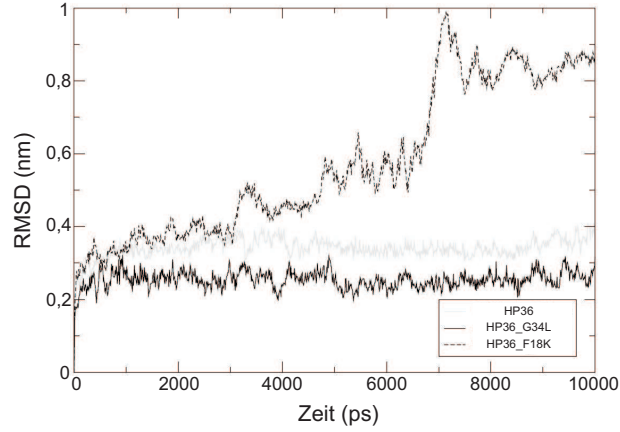
über die  $C_\alpha$ -Atome jeder Aminosäure des Peptidrückgrats berechnet, mit der Struktur als Referenz, welche der Anfangsstruktur von new2 bzw. 1vii entspricht.

Abbildung 4.17 zeigt die RMSF pro Aminosäure für die Strukturen aus den zwei Trajektorien für new2 (blaue und rote Linie) und den HP36-Wildtyp (1vii, schwarze Linie). Je höher der RMSF-Wert, desto flexibler sind die jeweiligen Bereiche des Peptids. Sehr hohe RMSF-Werte zwischen 0,3 und 0,75 nm lassen sich zwischen Aminosäuren 1 und 5 sowie 32 und 36 erkennen. Diese Bereiche überlappen mit den experimentell bestimmten flexiblen Termini des Peptids ( $\eta = -0,06$  bis  $-0,4$ ). Einen Anstieg des RMSF-Wertes auf 0,35 bzw. 0,5 nm lässt sich auch zwischen Aminosäuren 10 und 17 (new2) feststellen. Die hetNOE-Werte  $\eta$  zeigen einen weiteren, etwas weniger flexiblen Teil zwischen den Aminosäuren 14 und 20 ( $\eta = 0$ ).

Für den HP36-Wildtyp (1vii) ist bekannt, dass es sich dabei um eine strukturell stabile Domäne handelt, der Wildtyp somit erwartungsgemäß nur eine geringe Eigendynamik besitzt. Obwohl die RMSF-Werte des Wildtyps 1vii für den C-Terminus mit denen von new2 übereinstimmen, kann insgesamt von einer höheren Gesamtdynamik von new2 als von 1vii gesprochen werden. So liegen die Werte von new2 für die Ergebnisse der ersten Trajektorie (rote Linie) um das 2,4fache höher und die Werte von new2 aus der zweiten Trajektorie (blaue Linie) um das 3,4fache höher als die des Wildtyps 1vii. Während die Flexibilität des C-Terminus mit einer RMSF zwischen 0,5 und 0,68 nm zu der des Wildtyps bei 0,7 nm als fast gleich beschrieben werden kann, weichen die RMSF-Werte des N-Terminus deutlich von denen des Wildtyps ab. So liegt die RMSF bei etwa 0,32 nm für 1vii und bei 0,48 bis 0,75 nm für new2. Dies spricht somit für eine starke Dynamik im N-terminalen Bereich, die um das 2,3fache höher ist als die des Wildtyps. Der Anstieg der RMSF-Werte bei Aminosäuren 8 bis 14 (1vii) mit 0,23 nm bzw. 10 bis 18 (new2) mit 0,32 bis 0,5 nm lässt sich auf den flexibleren Mittelteil des Peptids zurückführen, welcher nicht durch eine Helix charakterisiert wird (vgl. Kap. 3.1.2).



Die RMSF wird über Gleichung 4.2 definiert und ist proportional zur RMSD (*Root Mean Square Deviation*) [136]. Wie bereits erwähnt, wird bei der RMSF die Standardabweichung einer Atom- zu einer Referenzposition für jede einzelne Aminosäure, gemittelt über die Zeit, berechnet. Die RMSD dagegen ermittelt die Standardabweichung einer Atom- zu einer Referenzposition für jeden Zeitpunkt, gemittelt über alle betrachteten Atome.



**Abbildung 4.18.:** RMSD-Werte über 10 ns MD-Simulation von drei Peptiden: dem Villin Headpiece Wildtyp (HP36), der instabilen L18K-Mutante und der stabilen G34L-Mutante (siehe auch 1.2). Verändert nach [27].

Kuzmanic *et al.*, 2010 beschreiben die RMSF als quadratisches Mittel über Gleichung 4.2 durch Einsetzen von Gleichung 4.3. Dabei steht  $N_S$  für die Anzahl der Strukturen im betrachteten Ensemble,  $N_a$  ist die Anzahl der Atome einer Struktur und der Vektor  $\vec{r}$  gibt die räumlichen Koordinaten des gegebenen Atoms an.  $\langle \vec{r}_k \rangle$  ist die durchschnittliche Position von Atom  $k$  über  $N_S$ .

$$\sqrt{\langle RMSF^2 \rangle} = \sqrt{\frac{1}{N_a} \sum_{i=1}^{N_a} RMSF_k^2} \quad (4.2)$$

$$RMSF_k = \sqrt{\frac{1}{N_S} \sum_{i=1}^{N_S} \|\vec{r}_{ik} - \langle \vec{r}_k \rangle\|^2} \quad (4.3)$$

Die RMSD definiert Kuzmanic *et al.*, 2010 als quadratisches Mittel über Gleichung 4.4 durch Einsetzen von Gleichung 4.5.  $S_i$  und  $S_j$  stehen für die beiden überlagerten Makromoleküle, zwischen deren Atomen die RMSD berechnet werden soll, Indizes  $i$  und  $j$  beziehen sich auf verschiedene Strukturen. Die RMSD (Gleichung 4.5) wird als ein Minimum ( $\min_{rot,trans}$ ) über alle Rotationen und Translationen einer der Strukturen, die mit der Referenzstruktur verglichen wird, berechnet.

$$\sqrt{\langle RMSD^2 \rangle} = \sqrt{\frac{2}{N_S(N_S - 1)} \sum_{i=1}^{N_S-1} \sum_{j>1}^{N_S} RMSD^2(S_i, S_j)} \quad (4.4)$$

$$RMSD(S_i, S_j) = \min \sqrt{\frac{1}{N_a} \sum_{k=1}^{N_a} \|\vec{r}_{ik} - \vec{r}_{jk}\|^2} \quad (4.5)$$

*rot,trans*

Die RMSF und die RMSD werden über Gleichung 4.6 in Bezug gesetzt und sind proportional zueinander.

$$\sqrt{\langle RMSD^2 \rangle} = \sqrt{\frac{2N_S}{N_S - 1} \langle RMSF^2 \rangle} \Rightarrow RMSD \propto RMSF \quad (4.6)$$

Die Berechnung der RMSD dient somit dazu, wie auch die Berechnung der RMSF, dynamische Eigenschaften eines Peptids zu bestimmen oder vorherzusagen [28, 137].

So beschreiben Hohm *et al.*, 2005, wie die RMSD bei der instabilen F18K-Mutante des Villin Headpiece (HP36) über den Zeitraum von 10 ns von 0,2 nm auf bis zu 1 nm ansteigt, wohingegen die RMSD für den Wildtyp während der gesamten Simulationszeit bei 0,25 nm bleibt (siehe Abb. 4.18). Eine solche Simulation würde für mut8 ähnliche Resultate hervorbringen, wie für die F18K-Mutante von HP36.

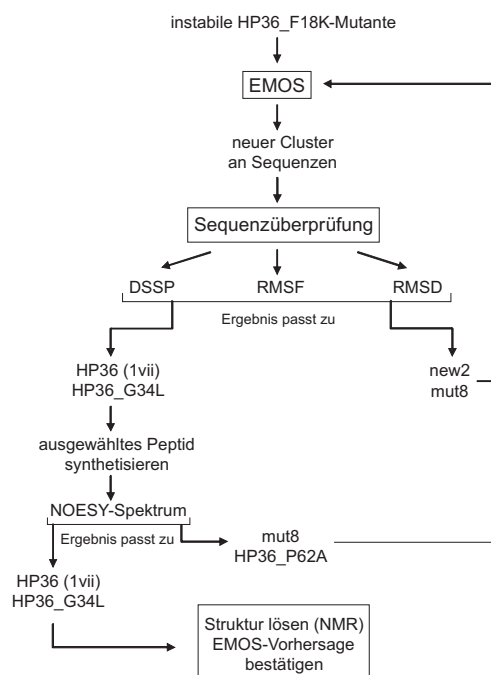
## 5. Ausblick

Die Resultate aus NMR-spektroskopischer Charakterisierung von mut8 und bioinformatischer Auswertung von new2 ergaben, dass es sich bei r1g11i1 um ein dynamisches Peptid handeln muss. Dessen Sekundärstrukturelemente sind zwar gut definiert, die Tertiärstruktur konnte jedoch nicht eindeutig bestimmt werden. Der evolutionäre Algorithmus hat somit das Kriterium Stabilität für r1g11i1 nicht eindeutig erfüllt und das Peptid als strukturell stabil deklariert. Da EMOS jedoch bereits erfolgreich an der stabilen G34L-Mutante von HP36 getestet wurde und somit Potential zeigt, sollte dieser Algorithmus verbessert werden. Nach einer Optimierung von EMOS werden weitere Durchläufe nötig sein, d. h. die Resultate einer neuen Voraussage müssen auch dann wieder überprüft werden. Da EMOS sehr komplexe Aufgaben erfüllen muss, werden auch mit großer Wahrscheinlichkeit noch mehrere Optimierungsläufe durchlaufen und zahlreiche Vorhersagen erstellt werden, die schnell und effizient untersucht werden sollten.

Diese Arbeit liefert eine Möglichkeit, Ergebnisse nach neuen EMOS-Durchläufen schneller und präziser zu analysieren. Durch die Kombination aus den hier beschriebenen Methoden (Berechnung der RMSF, RMSD, sowohl aus MD-Simulationsdaten und aus NMR-Daten und der direkte Bezug zueinander) wird eine neue, schnellere Methode zur Analyse von Resultaten gegeben, welche EMOS erstellt. Abbildung 5.1 zeigt, wie ein solcher Analyseweg aussehen kann. Gestartet wird einer neuer EMOS-Durchlauf mit der instabilen F18K-Mutante von Villin Headpiece 36 (HP36). Aus dem neuen Cluster an Sequenzen können dann diejenigen selektiert werden, welche bzgl. der Vorhersage (strukturelle Stabilität) als geeignete Kandidaten in engeren Betracht gezogen werden. Für diese werden wiederum über eine MD-Simulation Strukturen vorhergesagt und eine ausgewählt.

Daraufhin erfolgt eine Überprüfung der aus der MD-Simulation selektierten Struktur, und zwar weiterhin über bioinformatische Methoden. Nach einer weiteren MD-Simulation über eine bestimmte Zeit  $t$  bei einer Temperatur  $T$  von 300 K können die so ermittelten Strukturen der Trajektorie über DSSP ausgewertet werden (siehe Kapitel 4.1.2). Über DSSP würde sich eine Sekundärstrukturverteilung ergeben, die zum einen mit der des HP36-Wildtyps verglichen werden kann (drei präzise definierte  $\alpha$ -Helices von Aminosäuren 4 bis 11, 15 bis 20 und 23 bis 34 für 1vii) und zum anderen mit der von new2 bzw. new8 (zwei gut definierte  $\alpha$ -Helices von Aminosäuren 4 bis 8 und 27 bis 33 [gemittelte Werte aus bioinformatischer Vorhersage und NMR-Daten]). Die Voraussetzung dafür ist, dass eine gut definierte  $\alpha$ -Helix zu 80-100 % Sekundärstruktur-Anteil an den jeweiligen Sequenzpositionen vorliegen muss, um mit dem Wildtyp übereinzustimmen.

Die RMSF- und RMSD-Rechnungen liefern die präzisesten Aussagen über die vorliegende neue Vorhersage (siehe Kapitel 4.4). So können die Ergebnisse beider Aus-



**Abbildung 5.1.:** Analysemethode für Sequenzen nach einem neuen EMOS-Durchlauf.

wertungen (sowohl RMSF als auch RMSD) direkt mit den bereits bekannten RMSF- und RMSD-Daten über den HP36-Wildtyp, die G34L-Mutante, new2 und die F18K-Mutante verglichen werden. RMSF-Werte, die dabei – für nicht terminale Bereiche – durchschnittlich um 0,15 (und niedriger) liegen, deuten auf ein strukturell stabiles Peptid hin, RMSF-Werte um 0,25 (und höher) dagegen sprechen für ein eher instabiles Peptid, welches sich sehr schnell bewegt. RMSD-Werte sind für die G34L-Mutante, den HP36-Wildtyp und die F18K-Mutante vorhanden (siehe Abb. 4.18) und zeigen deutlich, dass für ein stabiles Peptid konstante Werte um 0,3 nm erwartet würden, während bei einem instabilen Peptid die RMSD nach 6 ns Simulationszeit von 0,3 auf 0,5 nm ansteigt. Voraussetzung ist hierbei eine Gesamt-Simulationszeit von min. 10 ns. Passen hier die über DSSP, RMSD und RMSF ermittelten Ergebnisse zu denen von mut8 bzw. new2, so handelt es sich sehr wahrscheinlich um ein instabiles Peptid und der Algorithmus muss weiter optimiert werden. Der Auswerteprozess würde an dieser Stelle enden, ohne dass die Mutante chemisch synthetisiert und über NMR-Spektroskopie charakterisiert werden muss, ein sowohl finanzieller als auch zeitlicher Vorteil. Lassen sich die Ergebnisse mit denen des HP36-Wildtyps oder der G34L-Mutante vergleichen, dann liegt die Wahrscheinlichkeit bei 80-90 %, dass das Peptid stabil ist.

Bei positiver Aussage kann zum nächsten Schritt übergegangen werden, das Peptid chemisch zu synthetisieren. Nach der Synthese wird ein NOESY-Spektrum aufgenommen und mit denen des HP36-Wildtyps, der stabilen G34L-Mutante, der instabilen P62A-Mutante und des mut8 verglichen. Sind in der Amidregion lediglich 25 bis 35 (+/-5) Signale zu sehen, so handelt es sich um ein instabiles Peptid. Sind mehr als 40

Signale zählbar, so liegt die Wahrscheinlichkeit bei 80-90 %, dass es sich um eine neue, strukturell stabile Mutante handelt. Im ersten Fall wird wiederum eine Optimierung von EMOS angestrebt, im zweiten Fall kann dann dazu übergegangen werden, die Struktur der neuen Mutante mittels NMR-Spektroskopie zu lösen. Interessanterweise wird in der Literatur beschrieben, dass der RMSF aus chemischen Verschiebungen (NMR) über den *Random Coil Index* vorhergesagt werden kann [138, 139, 140]. Diese Informationen können dann zusätzlich dazu dienen, die neue Mutante zu charakterisieren. Ebenfalls von Nutzen sein kann das HET-SOFAST-HMQC (siehe Kapitel 4.4), welches als 1D-Experiment kein  $^{15}\text{N}$ -markiertes Peptid benötigt und trotzdem Informationen zur Dynamik eines Peptids liefert. Wurde von der neuen Mutante ein Expressionskonstrukt hergestellt, so kann das Peptid auch  $^{15}\text{N}$ -markiert werden und ein hetNOE aufgenommen werden. Liegen die Werte bei etwa 0,2, so handelt es sich um ein instabiles, dynamisches Peptid, die vorangehenden Analysen waren also nicht selektiv genug. Liegen die Werte bei etwa 0,6, so liegt das Peptid strukturell stabil vor. Der Auswerteprozess würde für den EMOS-Durchlauf an dieser Stelle mit einer Positiv-Vorhersage enden.

## 6. Zusammenfassung

Villin Headpiece 36 (HP36) stellt als Teil des Villin eine autonom faltende kleine Untereinheit aus 36 Aminosäuren dar, dessen Tertiärstruktur 1997 von McKnight *et al.* über NMR-Spektroskopie gelöst wurde [17]. Aufgrund der geringen Größe und der sowohl strukturell als auch funktionell guten Charakterisierung, ist HP36 ein beliebtes Modellpeptid in der Bioinformatik. Diese nutzt HP36 in ausgestreckter Konformation, um die Faltung des Peptids zu dessen stabiler Tertiärstruktur über Zeiträume um 200 ns zu simulieren. Nach HP36 wurden auch Mutanten und Verwandte des Peptids auf deren Stabilität und Faltungseigenschaften untersucht. So wurde deutlich, dass einzelne Mutationen wie F18K oder P62A drastische Auswirkungen auf die Struktur des Peptids haben und dazu führen, dass HP36 nicht mehr nativ falten kann. Andere Austausche von Aminosäuren dagegen beeinflussen die Faltung kaum, so weicht die Tertiärstruktur der G34L-Mutante kaum von der des Wildtyps ab. Neben den Faltungssimulationen für HP36, ist die Bioinformatik auch daran interessiert, Peptide auf verschiedene Kriterien hin zu optimieren. Für die medizinische Forschung werden als Medikamente eingesetzte Peptide (z. B. Insulin) immer interessanter. So wäre es ein großer Fortschritt, wenn über bioinformatische Simulationen Peptide so designed werden könnten, dass sie z. B. resistent gegenüber Proteasen, gut löslich in wässriger Umgebung und nicht toxisch für den Menschen sind und zudem noch ihren medizinischen Zweck erfüllen. Solche bioinformatisch verfolgbaren Ziele würden dazu beitragen, wesentlich schneller, präziser und kostengünstiger Therapeutika auf Peptid-Basis herzustellen. Um diese Idee umsetzen zu können, werden hochkomplizierte Algorithmen benötigt, die mehrere Kriterien gleichzeitig optimieren können. Zu den bekannten Optimierungsverfahren in der Bioinformatik zählen evolutionäre Algorithmen, welche die natürliche Evolution mit Mutation und Selektion als Vorbild nehmen.

Der von der Bioinformatik-Gruppe um Prof. Dr. D. Hoffmann (ZMB, Universität Duisburg-Essen) entwickelte EMOS (*Evolutionary Multiobjective Optimization System*) ist ein evolutionärer Algorithmus, der auf mehrere Kriterien hin Peptide optimiert. Dieser Algorithmus wurde erfolgreich auf das Kriterium Stabilität hin getestet. Dafür wurde die instabile F18K-HP36-Mutante über EMOS wieder strukturell stabilisiert. Es ergab sich ein Ensemble an Strukturen, in dem u. a. auch die G34L-Mutante auftrat. Diese erwies sich nachweislich sogar als stabiler gegenüber dem Wildtyp, das von der Bioinformatik postulierte Ergebnis wurde über NMR-Spektroskopie verifiziert.

Eine andere Sequenz aus dem vorhergesagten Ensemble, r1g11i1, fiel dabei durch die hohe Anzahl an acht Mutationen auf. Für diese Mutante wurde von EMOS ebenfalls eine hohe strukturelle Stabilität vorhergesagt. Für r1g11i1 wurde über eine MD-Simulation die Struktur new2 vorhergesagt.

In dieser Arbeit wurde das chemisch synthetisierte r1g11i1, bezeichnet als mut8, über NMR-Spektroskopie charakterisiert und mit der EMOS-Vorhersage verglichen.

Dafür wurden homonukleare 2D-Spektren (COSY, TOCSY, NOESY) aufgenommen und zugeordnet. Die Informationen aus den Signalen des NOESY-Spektrums dienten dazu, sowohl eine Sekundärstrukturanalyse des Peptids durchzuführen als auch ein Ensemble an möglichen Tertiärstrukturen zu berechnen. Dabei wurde deutlich, dass mut8 gut definierte  $\alpha$ -Helices am N- und C-Terminus besitzt (Aminosäuren 5 bis 10 und 30 bis 34 [Mittelung der NMR-Daten]), wohingegen eine mittlere Helix (AS 23 bis 26) weniger gut bestimmt werden konnte. Entgegen der präzise bestimmbaren Sekundärstrukturen von mut8, konnte keine eindeutige Tertiärstruktur berechnet werden. mut8 aggregierte nicht in Lösung, einige NOEs waren nicht miteinander kombinierbar und trotzdem waren gut definierbare  $\alpha$ -Helices bestimmbar. Aufgrund dessen wurde vermutet, dass es sich bei der HP36-Mutante um ein dynamisches Peptid handelt, dessen Konformationen so schnell ineinander übergehen, dass sie nicht mehr über ein NOESY-Spektrum auflösbar sind. Um dieser Annahme nachkommen zu können, sollte ein heteronukleares HSQCNOE-Spektrum aufgenommen werden, wofür wiederum mut8  $^{15}\text{N}$ -markiert vorliegen musste. Dafür wurde mut8 erfolgreich kloniert,  $^{15}\text{N}$ -markiert exprimiert und gereinigt. Der N-terminale Abbau von mut8 während der Expression wurde zwar nicht gestoppt, jedoch minimiert. Es wurde ein HSQCNOE-Spektrum aufgenommen, dessen Signale mit Hilfe eines SOFAST-HMQC-Spektrums der chemisch synthetisierten mut8-Probe eindeutig zugeordnet werden konnten. Durch die Aufnahme des HSQCNOE-Spektrums und der Berechnung des hetNOE ( $\eta$ ) wurden die dynamischen Eigenschaften von mut8 bestätigt. So erwies sich mut8 als insgesamt dynamisch, besonders große Flexibilität konnte an den Termini sowie im mittleren Schleifen-Bereich des Peptids festgestellt werden.

Diesen Informationen zufolge hat EMOS das Ziel, mit r1g11i1 eine stabile Mutante vorherzusagen, nicht erreicht. Somit muss der Algorithmus weiterhin optimiert werden. Da r1g11i1 jedoch erstens eine stabile Sekundärstruktur aufweist und zweitens aus einem Cluster von Sequenzen stammt, für dessen G34L-Mutante eine erhöhte Stabilität nachgewiesen wurde, ist das Potential zur korrekten Vorhersage von struktureller Stabilität gegeben.

Die vorhergesagte Struktur new2 sollte abschließend darauf überprüft werden, ob sie sich genauso verhält wie mut8, d. h. ob dieselben Sekundärstrukturelemente zu finden und ob die dynamischen Eigenschaften ebenfalls gleich sind. Dafür wurden von Manuel Prinz (Bioinformatik, AG Hoffmann) über Bestimmung von NOEs und DSSP-Rechnungen gezeigt, dass in new2 die  $\alpha$ -Helices etwa von Aminosäuren 4 bis 10 und 23 bis 32 auftreten (eine mittlere Helix war auch hier nur mäßig gut definiert), diese Ergebnisse somit mit den ermittelten Werten für mut8 übereinstimmen. Ebenfalls durchgeführt wurden von Manuel Prinz RMSF-Kalkulationen, über welche die Dynamik von new2 bestätigt wurde. Auch hier stimmten die Daten mit denen für mut8 überein. Interessanterweise zeigte eine zusätzliche RMSF-Rechnung des HP36-Wildtyps (1vii), dass new2 bzw. mut8 sich insgesamt dynamischer verhält als der Wildtyp. Die Rechenleistung, die hinter EMOS steckt, ist äußerst komplex. So wird es auch weiterhin nötig sein – nach erfolgter Optimierung des Algorithmus – über EMOS getroffene Vorhersagen zu überprüfen. Die in dieser Arbeit gewonnenen Erkenntnisse können dazu dienen, zuerst EMOS zu verbessern und Resultate aus neuen Vorhersagen für HP36-Mutanten schnell und effizient zu überprüfen.

# Literaturverzeichnis

- [1] Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I. W., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. *Proteins* **50**(3), 437–450 Feb (2003).
- [2] Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. *Bioinformatics* **23**(21), 2947–2948 Nov (2007).
- [3] Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. *Bioinformatics* **25**(9), 1189–1191 May (2009).
- [4] Rehm, S. *Synthese und strukturelle Untersuchungen an Peptiden und Proteinen mit NMR- und Ultrakurzzeitspektroskopie*. PhD thesis, Fachbereich Biochemie, Chemie und Pharmazie der Johann Wolfgang Goethe - Universität in Frankfurt am Main, (2009).
- [5] Weicker, K. *Evolutionäre Algorithmen*. Teubner-Verlag, (2007).
- [6] Box, G. *Applied Statistics* **6**(2), 81–101 (1957).
- [7] Friedberg, R. M. *IBM Journal of Research and Development* **2**(1), 2–13 (1958).
- [8] Prinz, M. Master’s thesis, Fachhochschule Bingen, (2006).
- [9] Hohm, T., Limbourg, P., and Hoffmann, D. *J Comput Biol* **13**(1), 113–125 (2006).
- [10] Bretscher, A. and Weber, K. *Proc Natl Acad Sci U S A* **76**(5), 2321–2325 May (1979).
- [11] Glenney, J. R., Geisler, N., Kaulfus, P., and Weber, K. *J Biol Chem* **256**(15), 8156–8161 Aug (1981).
- [12] Bazari, W. L., Matsudaira, P., Wallek, M., Smeal, T., Jakes, R., and Ahmed, Y. *Proc Natl Acad Sci U S A* **85**(14), 4986–4990 Jul (1988).
- [13] Friederich, E., Vancompernelle, K., Louvard, D., and Vandekerckhove, J. *J Biol Chem* **274**(38), 26751–26760 Sep (1999).
- [14] Meng, J., Vardar, D., Wang, Y., Guo, H.-C., Head, J. F., and McKnight, C. J. *Biochemistry* **44**(36), 11963–11973 Sep (2005).



- [15] Zhai, L., Zhao, P., Panebra, A., Guerrierio, A. L., and Khurana, S. *J Biol Chem* **276**(39), 36163–36167 Sep (2001).
- [16] McKnight, C. J., Doering, D. S., Matsudaira, P. T., and Kim, P. S. *J Mol Biol* **260**(2), 126–134 Jul (1996).
- [17] McKnight, C. J., Matsudaira, P. T., and Kim, P. S. *Nat Struct Biol* **4**(3), 180–184 Mar (1997).
- [18] Arpin, M., Pringault, E., Finidori, J., Garcia, A., Jeltsch, J. M., Vandekerckhove, J., and Louvard, D. *J Cell Biol* **107**(5), 1759–1766 Nov (1988).
- [19] Hofmann, A., Noegel, A. A., Bomblies, L., Lottspeich, F., and Schleicher, M. *FEBS Lett* **328**(1-2), 71–76 Aug (1993).
- [20] Duan, Y., Wang, L., and Kollman, P. A. *Proc Natl Acad Sci U S A* **95**(17), 9897–9902 Aug (1998).
- [21] Lei, H., Wu, C., Liu, H., and Duan, Y. *Proc Natl Acad Sci U S A* **104**(12), 4925–4930 Mar (2007).
- [22] Lei, H., Deng, X., Wang, Z., and Duan, Y. *J Chem Phys* **129**(15), 155104 Oct (2008).
- [23] Piana, S., Laio, A., Marinelli, F., Troys, M. V., Bourry, D., Ampe, C., and Martins, J. C. *J Mol Biol* **375**(2), 460–470 Jan (2008).
- [24] Vermeulen, W., Vanhaesebrouck, P., Troys, M. V., Verschueren, M., Fant, F., Goethals, M., Ampe, C., Martins, J. C., and Borremans, F. A. M. *Protein Sci* **13**(5), 1276–1287 May (2004).
- [25] Vermeulen, W., Troys, M. V., Bourry, D., Dewitte, D., Rossenu, S., Goethals, M., Borremans, F. A. M., Vandekerckhove, J., Martins, J. C., and Ampe, C. *J Mol Biol* **359**(5), 1277–1292 Jun (2006).
- [26] Frank, B. S., Vardar, D., Buckley, D. A., and McKnight, C. J. *Protein Sci* **11**(3), 680–687 Mar (2002).
- [27] Hohm, T. and Hoffmann, D. *GECCO '05 Proceedings of the Genetic and Evolutionary Computation Conference*. **ACM 2005**, 423–429 (2005).
- [28] Mori, G. M. S. D., Colombo, G., and Micheletti, C. *Proteins* **58**(2), 459–471 Feb (2005).
- [29] Liwo, A., Khalili, M., and Scheraga, H. A. *Proc Natl Acad Sci U S A* **102**(7), 2362–2367 Feb (2005).
- [30] Rohl, C. A., Strauss, C. E. M., Chivian, D., and Baker, D. *Proteins* **55**(3), 656–677 May (2004).

- [31] Srinivasan, R. and Rose, G. D. *Proteins* **47**(4), 489–495 Jun (2002).
- [32] Böhm, H. J., Klebe, G., and Kubinyi, H. *Wirkstoffdesign*. Spektrum Akademischer Verlag, (1996).
- [33] Morrison, K. L. and Weiss, G. A. *Curr Opin Chem Biol* **5**(3), 302–307 Jun (2001).
- [34] Bäck, T. *Evolutionary Algorithms in Theory and Practice*. Oxford University Press, New York, (1996).
- [35] Deb, K. *Multi-Objective Optimization using Evolutionary Algorithms*. Wiley-Interscience Series in Systems and Optimization. John Wiley & Sons, Ltd., Baffins Lane, Chichester, West Sussex, England, (2001).
- [36] Gronwald, W., Hohm, T., and Hoffmann, D. *BMC Bioinformatics* **9**, 109 (2008).
- [37] Sanger Institute, Hinxton, Cambridge, UK. *Longden I: EMBOSS program prettyplot.*, (1999).
- [38] Lehrer, R. I. *Nat Rev Microbiol* **2**(9), 727–738 Sep (2004).
- [39] Pazdzierny, G. *junior RUBIN* (2003).
- [40] Zasloff, M. *Nature* **415**(6870), 389–395 Jan (2002).
- [41] Yagi, Y., Terada, K., Noma, T., Ikebukuro, K., and Sode, K. *BMC Bioinformatics* **8**, 11 (2007).
- [42] Hardy, H. and Skolnik, P. R. *Pharmacotherapy* **24**(2), 198–211 Feb (2004).
- [43] Watt, P. M. *Nat Biotechnol* **24**(2), 177–183 Feb (2006).
- [44] Farzan, M., Babcock, G. J., Vasilieva, N., Wright, P. L., Kiprilov, E., Mirzabekov, T., and Choe, H. *J Biol Chem* **277**(33), 29484–29489 Aug (2002).
- [45] Bannert, N., Craig, S., Farzan, M., Sogah, D., Santo, N. V., Choe, H., and Sodroski, J. *J Exp Med* **194**(11), 1661–1673 Dec (2001).
- [46] HI-Virus. *Abbott GmbH & Co. KG* (2010).
- [47] *Trofile and Tropism* (Monogram Biosciences, Inc. 2007).
- [48] Moore, J. P. and Doms, R. W. *Proc Natl Acad Sci U S A* **100**(19), 10598–10602 Sep (2003).
- [49] Kuritzkes, D. R., Jacobson, J., Powderly, W. G., Godofsky, E., DeJesus, E., Haas, F., Reimann, K. A., Larson, J. L., Yarbough, P. O., Curt, V., and Shanahan, W. R. *J Infect Dis* **189**(2), 286–291 Jan (2004).
- [50] Abel, S., Back, D. J., and Vourvahis, M. *Antivir Ther* **14**(5), 607–618 (2009).

- [51] Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., and Hendrickson, W. A. *Nature* **393**(6686), 648–659 Jun (1998).
- [52] Wang, J. H., Meijers, R., Xiong, Y., Liu, J. H., Sakihama, T., Zhang, R., Joachimiak, A., and Reinherz, E. L. *Proc Natl Acad Sci U S A* **98**(19), 10799–10804 Sep (2001).
- [53] Huang, C.-C., Lam, S. N., Acharya, P., Tang, M., Xiang, S.-H., Hussan, S. S.-U., Stanfield, R. L., Robinson, J., Sodroski, J., Wilson, I. A., Wyatt, R., Bewley, C. A., and Kwong, P. D. *Science* **317**(5846), 1930–1934 Sep (2007).
- [54] Dragic, T., Trkola, A., Thompson, D. A., Cormier, E. G., Kajumo, F. A., Maxwell, E., Lin, S. W., Ying, W., Smith, S. O., Sakmar, T. P., and Moore, J. P. *Proc Natl Acad Sci U S A* **97**(10), 5639–5644 May (2000).
- [55] Armand-Ugón, M., Moncunill, G., Gonzalez, E., Mena, M., Ballana, E., Clotet, B., and Esté, J. A. *J Antimicrob Chemother* **65**(3), 417–424 Mar (2010).
- [56] Seibert, C., Cadene, M., Sanfiz, A., Chait, B. T., and Sakmar, T. P. *Proc Natl Acad Sci U S A* **99**(17), 11031–11036 Aug (2002).
- [57] Cormier, E. G., Persuh, M., Thompson, D. A., Lin, S. W., Sakmar, T. P., Olson, W. C., and Dragic, T. *Proc Natl Acad Sci U S A* **97**(11), 5762–5767 May (2000).
- [58] Kühnemund, T. Master’s thesis, Institut für Organische Chemie, Hamburg., (2000).
- [59] Robertson, D. L., Hahn, B. H., and Sharp, P. M. *J Mol Evol* **40**(3), 249–259 Mar (1995).
- [60] Rambaut, A., Posada, D., Crandall, K. A., and Holmes, E. C. *Nat Rev Genet* **5**(1), 52–61 Jan (2004).
- [61] Marx, V. *CHEMICAL & Engineering News* **83**, 17–24 (2005).
- [62] Latham, P. W. *Nat Biotechnol* **17**(8), 755–757 Aug (1999).
- [63] Loffet, A. *J Pept Sci* **8**(1), 1–7 Jan (2002).
- [64] Bleicher, K. H., Böhm, H.-J., Müller, K., and Alanine, A. I. *Nat Rev Drug Discov* **2**(5), 369–378 May (2003).
- [65] Pritchard, J. F., Jurima-Romet, M., Reimer, M. L. J., Mortimer, E., Rolfe, B., and Cayen, M. N. *Nat Rev Drug Discov* **2**(7), 542–553 Jul (2003).
- [66] Scapin, G. *Curr Pharm Des* **12**(17), 2087–2097 (2006).
- [67] Hou, T. and Xu, X. *Curr Pharm Des* **10**(9), 1011–1033 (2004).
- [68] Jones, D. T. *Protein Sci* **3**(4), 567–574 Apr (1994).

- [69] Campbell, W., Kleiman, L., Barany, L., Li, Z., Khorchid, A., Fujita, E., Okada, N., and Okada, H. *Microbiol Immunol* **46**(3), 211–215 (2002).
- [70] Handschuh, S., Wagener, M., and Gasteiger, J. *J Chem Inf Comput Sci* **38**(2), 220–232 (1998).
- [71] Gillet, V. J., Willett, P., Fleming, P. J., and Green, D. V. S. *J Mol Graph Model* **20**(6), 491–498 Jun (2002).
- [72] Gillet, V. J., Khatib, W., Willett, P., Fleming, P. J., and Green, D. V. S. *J Chem Inf Comput Sci* **42**(2), 375–385 (2002).
- [73] Neumann, D. Master’s thesis, Universität Duisburg-Essen, ZMB, (2006).
- [74] Kleywegt, G. J., Zou, J., Kjeldgaard, M., and Jones, T. *International Tables for Crystallography, Chapter 17.1 Vol. F. Crystallography of Biological Macromolecules*, 353–356, 366–367 (2001).
- [75] Clamp, M., Cuff, J., Searle, S. M., and Barton, G. J. *Bioinformatics* **20**(3), 426–427 Feb (2004).
- [76] Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. *J Magn Reson* **160**(1), 65–73 Jan (2003).
- [77] Schwieters, C. D., Kuszewski, J. J., and Clore, G. M. *Progr NMR Spectroscopy* **48**, 47–62 (2006).
- [78] Gronwald, W. and Kalbitzer, H. *Progress in Nuclear Magnetic Resonance Spectroscopy* **44**, 33–96 (2004).
- [79] Goddard, T. D. and Kneller, D. G. University of California, San Francisco.
- [80] Bachem. *Bubendorf , Switzerland* (2006).
- [81] Marion, D., Ikura, M., Tschudin, R., and Bax, A. *Journal of Magnetic Resonance* **85**, 393–399 (1989).
- [82] Piotto, M., Saudek, V., and Sklenár, V. *J Biomol NMR* **2**(6), 661–665 Nov (1992).
- [83] Aue, W., Bartholdy, E., and Ernst, R. *Journal of Chemical Physics* **64**, 2229–2246 (1976).
- [84] Cavanagh, J. and Rance, M. *Journal of Magnetic Resonance* **88**(1), 72 – 85 (1990).
- [85] Jeener, J., Meier, B., Bachmann, P., and Ernst, R. *Journal of Chemical Physics* **71**, 4546–4553 (1979).
- [86] Schanda, P. and Brutscher, B. *J Am Chem Soc* **127**(22), 8014–8015 Jun (2005).

- [87] Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. *Biochemistry* **33**(19), 5984–6003 May (1994).
- [88] Kay, L. E., Torchia, D. A., and Bax, A. *Biochemistry* **28**(23), 8972–8979 Nov (1989).
- [89] Cavanagh, J., Fairbrother, W., Palmer, A., Skelton, N., and Rance, M. *Protein NMR Spectroscopy. Principles and Practice.* (2007).
- [90] Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. *J Biomol NMR* **5**(1), 67–81 Jan (1995).
- [91] Wüthrich, K. *NMR of Proteins and Nucleic Acids.* John Wiley & Sons, (1986).
- [92] Wagner, G., Neuhaus, D., Wörgötter, E., Vasák, M., Kägi, J. H., and Wüthrich, K. *J Mol Biol* **187**(1), 131–135 Jan (1986).
- [93] Wishart, D. S., Sykes, B. D., and Richards, F. M. *J Mol Biol* **222**(2), 311–333 Nov (1991).
- [94] Clore, G. M. and Gronenborn, A. M. *Proc Natl Acad Sci U S A* **95**(11), 5891–5898 May (1998).
- [95] Powell, M. *Mathematical Programming* **12**, 241–254 (1977).
- [96] RAMACHANDRAN, G. N., RAMAKRISHNAN, C., and SASISEKHARAN, V. *J Mol Biol* **7**, 95–99 Jul (1963).
- [97] Kami, K., Rajesh, S., and Overduin, M. *Methods Mol Biol* **462**, 291–306 (2009).
- [98] Muhle-Goll, D. C. and Grage, D. S. *Biochemie Vorlesung II, Einführung in die biomolekulare NMR-Spektroskopie, Lecture 4.* Forschungszentrum Karlsruhe.
- [99] Neudecker, P. B. *Strukturbestimmung von Birkenpollenallergenen und birkenpollenassoziierten Nahrungsmittelallergenen mit NMR-Spektroskopie.* PhD thesis, Universität Bayreuth, (2003).
- [100] Sander, C. *ProMoVal-Report, Computational Biology Center, Sander Research Lab .*
- [101] Wishart, D. S., Sykes, B. D., and Richards, F. M. *Biochemistry* **31**(6), 1647–1651 Feb (1992).
- [102] Cornilescu, G., Delaglio, F., and Bax, A. *J Biomol NMR* **13**(3), 289–302 Mar (1999).
- [103] Nilges, M., Kuszewski, J., and Brunger, A. *J C Hoch New York: Plenum Press* (1991).

- [104] Kubelka, J., Eaton, W. A., and Hofrichter, J. *J Mol Biol* **329**(4), 625–630 Jun (2003).
- [105] Neuhaus, D. and Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*. Wiley-Verlag, (2000).
- [106] Cho, H. S., Liu, C. W., Damberger, F. F., Pelton, J. G., Nelson, H. C., and Wemmer, D. E. *Protein Sci* **5**(2), 262–269 Feb (1996).
- [107] Williamson, M. P. *Modern Magnetic Resonance. The Nuclear Overhauser Effect*. Springer-Verlag, (2008).
- [108] Pfänder, R. *Kernspinresonanzspektroskopie an Proteinen in Lösung: Strukturbestimmung, Relaxationsstudien und dipolare Kopplungen*. PhD thesis, Max-Planck-Institut für Biochemie, Abteilung Strukturforschung, Biologische NMR-Spektroskopie, (2000).
- [109] Kabsch, W. and Sander, C. *Biopolymers* **22**(12), 2577–2637 Dec (1983).
- [110] Ng, C. A., Kato, Y., Tanokura, M., and Brownlee, R. T. C. *Biochim Biophys Acta* **1784**(9), 1208–1214 Sep (2008).
- [111] Pickart, C. M. and Eddins, M. J. *Biochim Biophys Acta* **1695**(1-3), 55–72 Nov (2004).
- [112] Dobson, N., Dantas, G., Baker, D., and Varani, G. *Structure* **14**(5), 847–856 May (2006).
- [113] Bader, R. and Zerbe, O. *Peptide NMR*. ETH Zürich.
- [114] Xiao, S., Bi, Y., Shan, B., and Raleigh, D. P. *Biochemistry* **48**(21), 4607–4616 Jun (2009).
- [115] Macias, M. J., Gervais, V., Civera, C., and Oschkinat, H. *Nat Struct Biol* **7**(5), 375–379 May (2000).
- [116] Jäger, M., Nguyen, H., Crane, J. C., Kelly, J. W., and Gruebele, M. *J Mol Biol* **311**(2), 373–393 Aug (2001).
- [117] Neidigh, J. W., Fesinmeyer, R. M., and Andersen, N. H. *Nat Struct Biol* **9**(6), 425–430 Jun (2002).
- [118] Neidigh, J. W., Fesinmeyer, R. M., Prickett, K. S., and Andersen, N. H. *Biochemistry* **40**(44), 13188–13200 Nov (2001).
- [119] Tisi, L. C. and Evans, P. A. *J Mol Biol* **249**(2), 251–258 Jun (1995).
- [120] Bunagan, M. R., Gao, J., Kelly, J. W., and Gai, F. *J Am Chem Soc* **131**(21), 7470–7476 Jun (2009).

- [121] Scott, K. A., Alonso, D. O. V., Sato, S., Fersht, A. R., and Daggett, V. *Proc Natl Acad Sci U S A* **104**(8), 2661–2666 Feb (2007).
- [122] Pervushin, K., Vögeli, B., and Eletsky, A. *J Am Chem Soc* **124**(43), 12898–12902 Oct (2002).
- [123] Brown, J. W., Vardar-Ulu, D., and McKnight, C. J. *J Mol Biol* **393**(3), 608–618 Oct (2009).
- [124] Frank, B. S., Vardar, D., Chishti, A. H., and McKnight, C. J. *J Biol Chem* **279**(9), 7909–7916 Feb (2004).
- [125] Markus, M. A., Dayie, K. T., Matsudaira, P., and Wagner, G. *Biochemistry* **35**(6), 1722–1732 Feb (1996).
- [126] Brüschweiler, R. *Curr Opin Struct Biol* **13**(2), 175–183 Apr (2003).
- [127] Wirmer, J., Peti, W., and Schwalbe, H. *J Biomol NMR* **35**(3), 175–186 Jul (2006).
- [128] Farrow, N. A., Zhang, O., Forman-Kay, J. D., and Kay, L. E. *Biochemistry* **36**(9), 2390–2402 Mar (1997).
- [129] Buevich, A. V., Shinde, U. P., Inouye, M., and Baum, J. *J Biomol NMR* **20**(3), 233–249 Jul (2001).
- [130] Dunker, A. K. and Obradovic, Z. *Nat Biotechnol* **19**(9), 805–806 Sep (2001).
- [131] Uversky, V. N. *Protein Sci* **11**(4), 739–756 Apr (2002).
- [132] Schanda, P., Forge, V., and Brutscher, B. *Magn Reson Chem* **44 Spec No**, S177–S184 Jul (2006).
- [133] Sousa, S. F., Fernandes, P. A., and Ramos, M. J. *Bioorg Med Chem* **17**(9), 3369–3378 May (2009).
- [134] Król, M., Roterman, I., Piekarska, B., Konieczny, L., Rybarska, J., Stopa, B., and Spólnik, P. *Proteins* **59**(3), 545–554 May (2005).
- [135] Yin, J., Bowen, D., and Southerland, W. M. *J Mol Graph Model* **24**(4), 233–243 Jan (2006).
- [136] Kuzmanic, A. and Zagrovic, B. *Biophys J* **98**(5), 861–871 Mar (2010).
- [137] Lee, I.-H., Kim, S.-Y., and Lee, J. *J Comput Chem* **31**(1), 57–65 Jan (2010).
- [138] Berjanskii, M. V. and Wishart, D. S. *J Biomol NMR* **40**(1), 31–48 Jan (2008).
- [139] Berjanskii, M. and Wishart, D. S. *Nat Protoc* **1**(2), 683–688 (2006).
- [140] Berjanskii, M. V. and Wishart, D. S. *Nucleic Acids Res* **35**(Web Server issue), W531–W537 Jul (2007).

# A. Protokolle

## generate.inp

```
remarks  file  nmr/generate.inp
remarks  Generate structure file for a protein
remarks  using the SA parameter and topology files.

topology
10  @/home/anja.matena/xplor/xplor-nih-2.18/toppar/topallhdg.pro

end                                     {*Read topology file          *}
parameter
  @/home/anja.matena/xplor/xplor-nih-2.18/toppar/parallhdg.pro

end
segment
  name="mut8"
20  chain                               {*Generate protein          *}
    @/home/anja.matena/xplor/xplor-nih-2.18/toppar/toph19.pep
    sequence
    MET TYR SER ALA GLU ASP PHE LYS ALA VAL
    ASP GLY MET THR ARG SER ALA LYS ALA ASN
    LEU PRO LEU GLU LYS GLN GLN ASN LEU LYS
    LYS ALA TYR TRP LEU PHE
  end                                   {*obtain the sequence        *}

end
30 end
write structure output=mut8.psf end
stop
```

## generate\_template.inp

```
remarks  file  nmr/generate_template.inp
remarks  Generates a "template" coordinate set.  This produces
remarks  an arbitrary extended conformation with ideal geometry.
remarks  Author: Axel T. Brunger

10  {====>}
    structure @mut8.psf end                                     {*Read structure file.*}

parameter
  {====>}
    @/home/anja.matena/xplor/xplor-nih-2.18/toppar/parallhdg.pro
    {*Read parameters.*}
end

20  vector ident (x) ( all )
    vector do (x=x/10.) ( all )
    vector do (y=random(0.5) ) ( all )
    vector do (z=random(0.5) ) ( all )

    vector do (fbeta=50) ( all )                                {*Friction coefficient, in 1/ps.*}
    vector do (mass=100) ( all )                                {*Heavy masses, in amus.*}
```



```
parameter
  nbonds
30      cutnb=5.5 rcon=20. nbxmod=-2 repel=0.9 wmin=0.1 tolerance=1.
      rexp=2 irexp=2 inhibit=0.25
  end
end

flags exclude * include bond angle vdw end

minimize powell nstep=50 nprint=10 end

flags include impr end
40 minimize powell nstep=50 nprint=10 end

dynamics verlet
  nstep=50 timestep=0.001 iasvel=maxwell firsttemp= 300.
  tcoupling = true tbath = 300. nprint=50 iprfreq=0
end

parameter
  nbonds
50      rcon=2. nbxmod=-3 repel=0.75
  end
end

minimize powell nstep=100 nprint=25 end

dynamics verlet
  nstep=500 timestep=0.005 iasvel=maxwell firsttemp= 300.
  tcoupling = true tbath = 300. nprint=100 iprfreq=0
end
60 flags exclude vdw elec end
vector do (mass=1.) ( name h* )
hbuild selection=( name h* ) phistep=360 end
flags include vdw elec end

minimize powell nstep=200 nprint=50 end

remarks produced by nmr/generate_template.inp          { *Write coordinates.* }
write coordinates output=mut8.pdb end
70 stop
```

### sa\_mut8.inp

```
remarks file nmr/sa.inp
remarks Simulated annealing protocol for NMR structure determination.
remarks The starting structure for this protocol can be any structure with
remarks a reasonable geometry, such as randomly assigned torsion angles or
remarks extended strands.
10 remarks Author: Michael Nilges

{====>}
evaluate ($init_t = 1000 )          { *Initial simulated annealing temperature.* }
{====>}
evaluate ($high_steps= 6000 )       { *Total number of steps at high temp.* }
{====>}
evaluate ($cool_steps = 3000 )      { *Total number of steps during cooling.* }

parameter                                { *Read the parameter file.* }
20 {====>}
    @/home/anja.matena/xplor/xplor-nih-2.18/toppar/parallhdg.pro
end
```

```

{====>}
structure @/home/anja.matena/md_mut8/mut8.psf end           {*Read the structure file
.*}

topology
@/home/anja.matena/xplor/xplor-nih-2.18/toppar/topallhdg.pro
end
30 {====>}
coordinates @/home/anja.matena/md_mut8/mut8.pdb           {*Read the coorndinates.*}

noe
{====>}
    nres=3000                {*Estimate greater than the actual number of NOEs.*}
    class seq
{====>}
40 @/home/anja.matena/md_mut8/constraints/NOE_constraints.tbl
end

!noe
!{====>}
!    nres=3000                {*Estimate greater than the actual number of NOEs.*}
!    class long
!{====>}
!end
50
    restraints
    dihedral
    reset
    @/home/anja.matena/md_mut8/constraints/TALOS_Constraints.tbl
end

flags exclude * include bonds angle impr vdw noe cdih end

                                {*Friction coefficient for MD heatbath, in 1/ps. *}
60 vector do (fbeta=10) (all)
                                {*Uniform heavy masses to speed molecular dynamics.*}
    vector do (mass=100) (all)

noe                                {*Parameters for NOE effective energy term.*}
    ceiling=1000
    averaging * cent
    potential * soft
    scale seq 50.
    scale long 10
70 sqoffset * 0.0
    sqconstant * 1.0
    sqexponent * 2
    soexponent * 1
    asymptote * 0.1                {*Initial value—modified later.*}
    rswitch * 0.5
end

parameter                        {*Parameters for the repulsive energy term.*}
    nbonds
80     repel=1.                    {*Initial value for repel—modified later.*}
        rexp=2 irexp=2 rcon=1.
        nbxmod=3
        wmin=0.01
        cutnb=4.5 ctonnb=2.99 ctofnb=3.
        tolerance=0.5
    end
end

restraints dihedral
90     scale=20.

```

```

end

{====>}
evaluate ($end_count=100)          {*Loop through a family of 100 structures.*}

coor copy end

evaluate ($count = 0)
while ($count < $end_count ) loop main
100   evaluate ($count=$count+1)

      coor swap end
      coor copy end

      {* ===== Initial minimization.*}
      restraints dihedral   scale=40.   end
      noe asymptote * 0.1   end
      parameter nbonds repel=1.   end end
110   constraints interaction
          (not name SG) (all) weights * 1 vdw 0.002 end end
      minimize powell nstep=50 drop=10. nprint=25 end

      {* ===== High-temperature dynamics.*}
      constraints interaction (not name SG) (all)
          weights * 1 angl 0.4 impr 0.1 vdw 0.002 end end

      evaluate ($nstep1=int($high_steps * 2. / 3. ) )
120   evaluate ($nstep2=int($high_steps * 1. / 3. ) )

      dynamics verlet
          nstep=$nstep1 timestep=0.002 iasvel=maxwell firstt=$init_t
          tcoupling=true tbath=$init_t nprint=50 iprfreq=0
      end

      {* ===== Tilt the asymptote and increase weights on geometry.*}
      noe asymptote * 1.0   end
130   constraints interaction
          (not name SG) (all) weights * 1 vdw 0.002 end end
      dynamics verlet
          nstep=$nstep2 timestep=0.002 iasvel=current tcoupling=true
          tbath=$init_t nprint=50 iprfreq=0
      end

      {* ===== Cool the system.*}

140   restraints dihedral   scale=200.   end

      evaluate ($final_t = 100)          { K }
      evaluate ($tempstep = 50)          { K }

      evaluate ($ncycle = ($init_t-$final_t)/$tempstep)
      evaluate ($nstep = int($cool_steps/$ncycle))

      evaluate ($ini_rad = 0.9)           evaluate ($fin_rad = 0.75)
      evaluate ($ini_con= 0.003)          evaluate ($fin_con= 4.0)

150   evaluate ($bath = $init_t)
      evaluate ($k_vdw = $ini_con)
      evaluate ($k_vdwfact = ($fin_con/$ini_con)^(1/$ncycle))
      evaluate ($radius= $ini_rad)
      evaluate ($radfact = ($fin_rad/$ini_rad)^(1/$ncycle))

      evaluate ($i_cool = 0)
      while ($i_cool < $ncycle) loop cool

```

```

160     evaluate ($i_cool=$i_cool+1)

    evaluate ($bath = $bath - $tempstep)
    evaluate ($k_vdw=min($fin_con,$k_vdw*$k_vdwfact))
    evaluate ($radius=max($fin_rad,$radius*$radfact))

    parameter nbonds repel=$radius end end
    constraints interaction (not name SG) (all)
        weights * 1. vdw $k_vdw end end

    dynamics verlet
170     nstep=$nstep time=0.002 iasvel=current firstt=$bath
        tcoup=true tbath=$bath nprint=$nstep iprfreq=0
    end

    {====>}                                     {*Abort condition.*}
        evaluate ($critical=$temp/$bath)
        if ($critical > 10. ) then
            display ****&&&& rerun job with smaller timestep (i.e., 0.003)
            stop
180     end if

end loop cool

{* ===== Final minimization.*}

constraints interaction (all) (all) weights * 1. vdw 4. end end
minimize powell nstep=200 drop=10.0 nprint=25 end

{* ===== Write out the final structure(s).*}
190 print threshold=0.2 noe
    evaluate ($rms_noe=$result)
    evaluate ($violations_noe=$violations)
    print threshold=5. cdih
    evaluate ($rms_cdih=$result)
    evaluate ($violations_cdih=$violations)
    print thres=0.05 bonds
    evaluate ($rms_bonds=$result)
    print thres=5. angles
    evaluate ($rms_angles=$result)
200    print thres=5. impropers
    evaluate ($rms_impropers=$result)
    remarks =====
    remarks          overall,bonds,angles,improper,vdw,noe,cdih
    remarks energies: $ener, $bond, $angl, $impr, $vdw, $noe, $cdih
    remarks =====
    remarks          bonds,angles,impropers,noe,cdih
    remarks rms-d: $rms_bonds,$rms_angles,$rms_impropers,$rms_noe,$rms_cdih
    remarks =====
    remarks          noe, cdih
210    remarks violations.: $violations_noe, $violations_cdih
    remarks =====

    {====>}                                     {*Name(s) of the family of final structures.*}
    evaluate ($filename="/home/anja.matena/md_mut8/SA_mut8/sa_mut8_"+encode($count)+".
        pdb")

    write coordinates output =$filename end

    remarks output file for violations only
    evaluate ($outfile="/home/anja.matena/md_mut8/SA_mut8/data_sal.out")
220    set display=$outfile end
    display structure= $filename
    display energies: $ener, $bond, $angl, $impr, $vdw, $noe, $cdih
    display rms-d: $rms_bonds, $rms_angles, $rms_impropers, $rms_noe, $rms_cdih
    display violations: $violations_noe, $violations_cdih
    set print=$outfile end

```

```

display *****
print threshold=0.2 noe
display *****
print threshold=5. cdih
230 display *****
print thres=0.05 bonds
display *****
print thres=5. angles
display *****
print thres=5. impropers
display
*****

set print=schrott_sa end

240 end loop main

stop

ref_mut8.inp

remarks file nmr/refine.inp — Simulated annealing refinement
remarks for NMR structure determination
remarks
remarks Authors: Michael Nilges, John Kuszewski, and Axel T. Brunger

10 !
! Please cite the following paper when using this protocol:
! 1. Nilges, M., Kuszewski, J., and Brunger, A.T. (1991).
! In: Computational Aspects of the Study of Biological
! Macromolecules by NMR, (J.C. Hoch, ed.). New York: Plenum Press.
!

{====>}
evaluate ($init_t = 1000)          {*Initial annealing temperature, in K.*}
{====>}
20 evaluate ($cool_steps = 4000 )  {*Total number of steps during cooling.*}

parameter                         {*Read the parameter file.*}
{====>}
@/home/anja.matena/xplor/xplor-nih-2.18/toppar/parallhdg.pro
end

{====>}
structure @/home/anja.matena/md_mut8/mut8.psf end {*The structure file.*}

30 topology
@/home/anja.matena/xplor/xplor-nih-2.18/toppar/topallhdg.pro
end

!patch CIPP
!reference===(resid 113)
!reference+==(resid 114)
!end

40 noe
{====>}
nres=3000                        {*Estimate greater than the actual number of NOEs.*}
class all
{====>}

@/home/anja.matena/md_mut8/constraints/NOE_constraints.tbl
end

{====>}

```

```

50 restraints
    dihedral
    reset
    @/home/anja.matena/md_mut8/constraints/TALOS_constraints.tbl
end

                                { *Friction coefficient for MD heatbath, in 1/ps.*}
vector do (fbeta=10) (all)
vector do (mass=100) (all)      { *Heavy masses to speed molecular dynamics.*}

60 noe                          { *Parameters for NOE effective energy term.*}
    ceiling=1000
    averaging * cent
    potential * square
    sqconstant * 1.
    sqexponent * 2
    scale * 50.                { *Constant NOE scale throughout the protocol.*}
end

parameter                      { *Parameters for the repulsive energy term.*}
70 nbonds
    repel=0.5                  { *Initial value for repel—modified later.*}
    rexp=2 irexp=2 rcon=1.
    nbxmod=3
    wmin=0.01
    cutnb=4.5 ctonnb=2.99 ctofnb=3.
    tolerance=0.5
end

80 restraints dihedral
    scale=200.
end

{====>}
evaluate ($end_count=100)      { *Loop through a family of 100 structures.*}

evaluate ($count = 0)
while ($count < $end_count ) loop main

90     evaluate ($count=$count+1)

{====>}                        { *Filename(s) for embedded coordinates.*}
    evaluate ($filename="/home/anja.matena/md_mut8/SA_mut8/sa_mut8_"+encode($count)+".
        pdb")

    coor @@ $filename

    flags exclude * include bond angl impr vdw noe cdih end

    vector do (vx=maxwell($init_t)) ( all )
    vector do (vy=maxwell($init_t)) ( all )
100    vector do (vz=maxwell($init_t)) ( all )

    evaluate ($final_t = 100)    { K }
    evaluate ($tempstep = 50)   { K }

    evaluate ($ncycle = ($init_t-$final_t)/$tempstep)
    evaluate ($nstep = int($cool_steps/$ncycle))

    evaluate ($ini_rad = 0.9)      evaluate ($fin_rad = 0.75)
110    evaluate ($ini_con= 0.003)   evaluate ($fin_con= 4.0)

    evaluate ($bath = $init_t)
    evaluate ($k_vdw = $ini_con)
    evaluate ($k_vdwfact = ($fin_con/$ini_con)^(1/$ncycle))
    evaluate ($radius= $ini_rad)
    evaluate ($radfact = ($fin_rad/$ini_rad)^(1/$ncycle))

```

```

evaluate ($i_cool = 0)
while ($i_cool < $ncycle) loop cool
120   evaluate ($i_cool=$i_cool+1)

   evaluate ($bath = $bath - $tempstep)
   evaluate ($k_vdw=min($fin_con,$k_vdw*$k_vdwfact))
   evaluate ($radius=max($fin_rad,$radius*$radfact))

   parameter nbonds repel=$radius end end
   constraints interaction (all) (all) weights * 1. vdw $k_vdw end end

   dynamics verlet
130     nstep=$nstep time=0.002 iasvel=current firstt=$bath
        tcoup=true tbath=$bath nprint=$nstep iprfreq=0
   end

{====>}                                     {*Abort condition.*}
   evaluate ($critical=$temp/$bath)
   if ($critical > 10. ) then
       display ****&&&& rerun job with smaller timestep (i.e., 0.003)
       stop
140   end if
end loop cool

{* ===== Final minimization.*}
minimize powell nstep= 200 nprint=25 end

{* ===== Write out the final structure(s).*}
print threshold=0.2 noe
evaluate ($rms_noe=$result)
evaluate ($violations_noe=$violations)
150 print threshold=5. cdih
evaluate ($rms_cdih=$result)
evaluate ($violations_cdih=$violations)
print thres=0.05 bonds
evaluate ($rms_bonds=$result)
print thres=5. angles
evaluate ($rms_angles=$result)
print thres=5. impropers
evaluate ($rms_impropers=$result)
remarks =====
160 remarks          overall , bonds , angles , improper , vdw , noe , cdih
remarks energies: $ener, $bond, $angl, $impr, $vdw, $noe, $cdih
remarks =====
remarks          bonds , angles , impropers , noe , cdih
remarks rms-d: $rms_bonds, $rms_angles, $rms_impropers, $rms_noe, $rms_cdih
remarks =====
remarks          noe, cdih
remarks violations.: $violations_noe, $violations_cdih
remarks =====

170 {====>}                                     {*Name(s) of the family of final structures.*}
   evaluate ($filename="/home/anja.matena/md_mut8/SA_mut8/refine1_"+encode($count)+".
       pdb")

write coordinates output =$filename end

remarks output file for violations only
evaluate ($outfile="/home/anja.matena/md_mut8/SA_mut8/refldata_sal.out")

set display=$outfile end
180 display structure= $filename
display energies: $ener, $bond, $angl, $impr, $vdw, $noe, $cdih
display rms-d: $rms_bonds, $rms_angles, $rms_impropers, $rms_noe, $rms_cdih
display violations: $violations_noe, $violations_cdih

```

```

    set print=$outfile end
    display *****
    print threshold=0.2 noe
    display *****
    print threshold=5. cdih
    display *****
190  print thres=0.05 bonds
    display *****
    print thres=5. angles
    display *****
    print thres=5. impropers
    display
        *****

    set print=schrott5_re end

end loop main

200 stop
```



## B. Ergebnisse

### NOE\_constraints.tbl

intraresidual											
	assign	(resid	2 and	name	HA)	(resid	2 and	name	HD*)	3.7	1.9 1.1
	assign	(resid	2 and	name	HD2)	(resid	2 and	name	HB1)	3.0	1.2 0.9
	assign	(resid	2 and	name	HE2)	(resid	2 and	name	HB1)	4.5	2.7 1.4
	assign	(resid	3 and	name	HB2)	(resid	3 and	name	HA)	3.9	2.1 1.2
10	assign	(resid	4 and	name	HA)	(resid	4 and	name	HN)	3.8	2.0 1.1
	assign	(resid	4 and	name	HB*)	(resid	4 and	name	HA)	3.0	1.2 0.9
	assign	(resid	4 and	name	HB*)	(resid	4 and	name	HN)	3.7	1.9 1.1
	assign	(resid	5 and	name	HA)	(resid	5 and	name	HB1)	3.7	1.9 1.1
	assign	(resid	5 and	name	HA)	(resid	5 and	name	HG2)	3.8	2.0 1.2
	assign	(resid	5 and	name	HB1)	(resid	5 and	name	HB2)	2.3	0.5 0.7
	assign	(resid	5 and	name	HB2)	(resid	5 and	name	HA)	4.0	2.2 1.2
	assign	(resid	5 and	name	HB2)	(resid	5 and	name	HN)	3.7	1.9 1.1
	assign	(resid	5 and	name	HG2)	(resid	5 and	name	HB1)	2.9	1.1 0.9
	assign	(resid	5 and	name	HG2)	(resid	5 and	name	HB2)	2.5	0.7 0.7
20	assign	(resid	5 and	name	HN)	(resid	5 and	name	HA)	4.4	2.6 1.3
	assign	(resid	5 and	name	HN)	(resid	5 and	name	HB1)	3.5	1.7 1.1
	assign	(resid	5 and	name	HN)	(resid	5 and	name	HG2)	4.4	2.6 1.3
	assign	(resid	6 and	name	HA)	(resid	6 and	name	HN)	3.4	1.6 1.0
	assign	(resid	6 and	name	HB2)	(resid	6 and	name	HA)	3.7	1.9 1.1
	assign	(resid	6 and	name	HN)	(resid	6 and	name	HB2)	3.2	1.4 1.0
	assign	(resid	7 and	name	HA)	(resid	7 and	name	HB1)	3.3	1.5 1.0
	assign	(resid	7 and	name	HA)	(resid	7 and	name	HB2)	3.4	1.6 1.0
	assign	(resid	7 and	name	HA)	(resid	7 and	name	HD*)	3.6	1.8 1.1
	assign	(resid	7 and	name	HB1)	(resid	7 and	name	HE*)	4.5	2.7 1.3
30	assign	(resid	7 and	name	HB2)	(resid	7 and	name	HB1)	2.1	0.3 0.6
	assign	(resid	7 and	name	HD*)	(resid	7 and	name	HB1)	3.5	1.7 1.1
	assign	(resid	7 and	name	HD*)	(resid	7 and	name	HB2)	3.5	1.7 1.0
	assign	(resid	7 and	name	HD*)	(resid	7 and	name	HN)	4.2	2.4 1.3
	assign	(resid	7 and	name	HN)	(resid	7 and	name	HB1)	3.4	1.6 1.0
	assign	(resid	7 and	name	HN)	(resid	7 and	name	HB2)	3.6	1.8 1.1
	assign	(resid	8 and	name	HA)	(resid	8 and	name	HN)	3.4	1.6 1.0
	assign	(resid	8 and	name	HB1)	(resid	8 and	name	HE2)	4.3	2.5 1.3
	assign	(resid	8 and	name	HB1)	(resid	8 and	name	HN)	3.1	1.3 0.9
	assign	(resid	8 and	name	HB2)	(resid	8 and	name	HA)	3.0	1.2 0.9
40	assign	(resid	8 and	name	HN)	(resid	8 and	name	HB2)	4.1	2.3 1.2
	assign	(resid	9 and	name	HA)	(resid	9 and	name	HB*)	2.4	0.6 0.7
	assign	(resid	9 and	name	HN)	(resid	9 and	name	HB*)	3.0	1.2 0.9
	assign	(resid	10 and	name	HB)	(resid	10 and	name	HA)	3.5	1.7 1.0
	assign	(resid	10 and	name	HB)	(resid	10 and	name	HN)	3.4	1.6 1.0
	assign	(resid	10 and	name	HG1*)	(resid	10 and	name	HA)	3.3	1.5 1.0
	assign	(resid	10 and	name	HG1*)	(resid	10 and	name	HB)	2.7	0.9 0.8
	assign	(resid	10 and	name	HG1*)	(resid	10 and	name	HN)	4.1	2.3 0.9
	assign	(resid	10 and	name	HN)	(resid	10 and	name	HA)	4.0	2.2 1.2
	assign	(resid	11 and	name	HA)	(resid	11 and	name	HB2)	3.0	1.2 0.9
50	assign	(resid	11 and	name	HA)	(resid	11 and	name	HN)	3.0	1.2 0.9
	assign	(resid	11 and	name	HB2)	(resid	11 and	name	HN)	3.0	1.2 0.9
	assign	(resid	12 and	name	HA1)	(resid	12 and	name	HN)	3.0	1.2 0.9
	assign	(resid	13 and	name	HA)	(resid	13 and	name	HB2)	3.0	1.2 0.9
	assign	(resid	13 and	name	HA)	(resid	13 and	name	HG1)	3.8	2.0 1.1
	assign	(resid	13 and	name	HB2)	(resid	13 and	name	HG1)	3.0	1.2 0.9
	assign	(resid	13 and	name	HB2)	(resid	13 and	name	HN)	3.0	1.2 0.9
	assign	(resid	13 and	name	HG1)	(resid	13 and	name	HN)	4.4	2.6 1.3
	assign	(resid	13 and	name	HG2)	(resid	13 and	name	HB2)	3.4	1.6 1.0
	assign	(resid	13 and	name	HG2)	(resid	13 and	name	HN)	4.4	2.6 1.3

60	assign	(resid	14	and	name	HG1)	(resid	14	and	name	HA)	2.9	1.1	0.9
	assign	(resid	14	and	name	HN)	(resid	14	and	name	HA)	3.4	1.6	1.0
	assign	(resid	15	and	name	HA)	(resid	15	and	name	HD2)	4.1	2.3	1.2
	assign	(resid	15	and	name	HB1)	(resid	15	and	name	HA)	4.1	2.3	1.2
	assign	(resid	15	and	name	HB1)	(resid	15	and	name	HN)	3.9	2.1	1.2
	assign	(resid	15	and	name	HB2)	(resid	15	and	name	HA)	3.9	2.1	1.2
	assign	(resid	15	and	name	HB2)	(resid	15	and	name	HB1)	2.1	0.3	0.6
	assign	(resid	15	and	name	HB2)	(resid	15	and	name	HN)	4.2	2.4	1.3
	assign	(resid	15	and	name	HD2)	(resid	15	and	name	HB1)	3.6	1.8	1.1
	assign	(resid	15	and	name	HD2)	(resid	15	and	name	HB2)	3.6	1.8	1.1
70	assign	(resid	15	and	name	HG2)	(resid	15	and	name	HA)	4.8	3.0	1.4
	assign	(resid	15	and	name	HG2)	(resid	15	and	name	HB2)	2.6	0.8	0.8
	assign	(resid	15	and	name	HG2)	(resid	15	and	name	HD2)	3.0	1.2	0.9
	assign	(resid	15	and	name	HG2)	(resid	15	and	name	HN)	4.1	2.3	1.2
	assign	(resid	16	and	name	HA)	(resid	16	and	name	HB2)	3.7	1.9	1.1
	assign	(resid	16	and	name	HA)	(resid	16	and	name	HN)	4.2	2.4	1.3
	assign	(resid	16	and	name	HB2)	(resid	16	and	name	HN)	4.1	2.3	1.2
	assign	(resid	17	and	name	HB*)	(resid	17	and	name	HA)	3.7	1.9	1.1
	assign	(resid	19	and	name	HB*)	(resid	19	and	name	HN)	3.2	1.4	0.9
	assign	(resid	19	and	name	HA)	(resid	19	and	name	HB*)	3.0	1.2	3.0
80	assign	(resid	20	and	name	HA)	(resid	20	and	name	HB1)	3.5	1.7	1.1
	assign	(resid	20	and	name	HA)	(resid	20	and	name	HB2)	3.0	1.2	0.9
	assign	(resid	20	and	name	HB1)	(resid	20	and	name	HB2)	2.0	0.2	0.6
	assign	(resid	20	and	name	HB2)	(resid	20	and	name	HN)	4.4	2.6	1.3
	assign	(resid	21	and	name	HA)	(resid	21	and	name	HB*)	3.6	1.8	1.1
	assign	(resid	21	and	name	HA)	(resid	21	and	name	HD1*)	4.2	2.4	1.0
	assign	(resid	21	and	name	HA)	(resid	21	and	name	HG)	3.2	1.4	1.0
	assign	(resid	21	and	name	HA)	(resid	21	and	name	HN)	3.5	1.7	1.0
	assign	(resid	21	and	name	HB*)	(resid	21	and	name	HG)	2.2	0.4	0.7
	assign	(resid	21	and	name	HD1*)	(resid	21	and	name	HG)	3.0	1.2	0.9
90	assign	(resid	21	and	name	HN)	(resid	21	and	name	HB*)	3.0	1.2	0.9
	assign	(resid	21	and	name	HN)	(resid	21	and	name	HD1*)	4.7	2.9	1.4
	assign	(resid	21	and	name	HN)	(resid	21	and	name	HG)	3.9	2.1	1.2
	assign	(resid	22	and	name	HB2)	(resid	22	and	name	HG*)	2.9	1.1	0.9
	assign	(resid	23	and	name	HA)	(resid	23	and	name	HN)	2.9	1.1	0.9
	assign	(resid	23	and	name	HB*)	(resid	23	and	name	HA)	3.5	1.7	1.1
	assign	(resid	23	and	name	HB*)	(resid	23	and	name	HD1*)	2.9	1.1	0.9
	assign	(resid	23	and	name	HB*)	(resid	23	and	name	HN)	2.9	1.1	0.9
	assign	(resid	23	and	name	HD1*)	(resid	23	and	name	HA)	3.6	1.8	1.1
	assign	(resid	23	and	name	HG)	(resid	23	and	name	HA)	3.4	1.6	1.0
100	assign	(resid	23	and	name	HN)	(resid	23	and	name	HD1*)	5.0	3.2	1.5
	assign	(resid	23	and	name	HN)	(resid	23	and	name	HG)	3.6	1.8	1.1
	assign	(resid	24	and	name	HA)	(resid	24	and	name	HB2)	2.8	1.0	0.9
	assign	(resid	24	and	name	HB1)	(resid	24	and	name	HA)	3.1	1.3	0.9
	assign	(resid	24	and	name	HB1)	(resid	24	and	name	HG2)	2.9	1.1	0.9
	assign	(resid	24	and	name	HB2)	(resid	24	and	name	HG2)	3.0	1.2	0.9
	assign	(resid	24	and	name	HG2)	(resid	24	and	name	HA)	3.8	2.0	1.1
	assign	(resid	24	and	name	HG2)	(resid	24	and	name	HB2)	2.4	0.6	0.7
	assign	(resid	24	and	name	HG2)	(resid	24	and	name	HN)	3.6	1.8	1.1
	assign	(resid	24	and	name	HN)	(resid	24	and	name	HB1)	3.4	1.6	1.0
110	assign	(resid	24	and	name	HN)	(resid	24	and	name	HB2)	3.5	1.7	1.1
	assign	(resid	25	and	name	HA)	(resid	25	and	name	HG2)	3.7	1.9	1.1
	assign	(resid	25	and	name	HB1)	(resid	25	and	name	HA)	3.7	1.9	1.1
	assign	(resid	25	and	name	HB1)	(resid	25	and	name	HN)	2.8	1.0	0.8
	assign	(resid	25	and	name	HB2)	(resid	25	and	name	HA)	3.6	1.8	1.1
	assign	(resid	25	and	name	HD2)	(resid	25	and	name	HE2)	3.1	1.3	0.9
	assign	(resid	25	and	name	HD2)	(resid	25	and	name	HG2)	2.3	0.5	0.7
	assign	(resid	25	and	name	HE2)	(resid	25	and	name	HG2)	4.1	2.3	1.2
	assign	(resid	25	and	name	HE2)	(resid	25	and	name	HN)	5.7	3.9	1.7
	assign	(resid	25	and	name	HG2)	(resid	25	and	name	HB2)	3.0	1.2	0.9
120	assign	(resid	25	and	name	HN)	(resid	25	and	name	HB2)	3.7	1.9	1.1
	assign	(resid	26	and	name	HA)	(resid	26	and	name	HG2)	3.1	1.3	0.9
	assign	(resid	26	and	name	HB1)	(resid	26	and	name	HA)	3.6	1.8	1.1
	assign	(resid	26	and	name	HB1)	(resid	26	and	name	HG2)	3.0	1.2	0.9
	assign	(resid	26	and	name	HB1)	(resid	26	and	name	HN)	3.4	1.6	1.0
	assign	(resid	26	and	name	HB2)	(resid	26	and	name	HA)	3.3	1.5	1.0
	assign	(resid	26	and	name	HB2)	(resid	26	and	name	HB1)	2.4	0.6	0.7
	assign	(resid	26	and	name	HB2)	(resid	26	and	name	HN)	3.9	2.1	1.2

	assign	(resid	26	and	name	HG2)	(resid	26	and	name	HB2)	2.5	0.7	0.8
	assign	(resid	26	and	name	HG2)	(resid	26	and	name	HN)	3.8	2.0	1.1
130	assign	(resid	27	and	name	HB1)	(resid	27	and	name	HB2)	2.3	0.5	0.7
	assign	(resid	27	and	name	HB1)	(resid	27	and	name	HN)	3.5	1.7	1.0
	assign	(resid	27	and	name	HB2)	(resid	27	and	name	HN)	3.8	2.0	1.2
	assign	(resid	27	and	name	HG2)	(resid	27	and	name	HB1)	2.7	0.9	0.8
	assign	(resid	27	and	name	HG2)	(resid	27	and	name	HB2)	2.9	1.1	0.9
	assign	(resid	27	and	name	HG2)	(resid	27	and	name	HN)	3.9	2.1	1.2
	assign	(resid	28	and	name	HB1)	(resid	28	and	name	HN)	4.5	2.7	1.3
	assign	(resid	28	and	name	HB2)	(resid	28	and	name	HB1)	2.2	0.4	0.7
	assign	(resid	28	and	name	HB2)	(resid	28	and	name	HN)	4.7	2.9	1.4
	assign	(resid	29	and	name	HA)	(resid	29	and	name	HD1*)	3.6	1.8	1.1
140	assign	(resid	29	and	name	HA)	(resid	29	and	name	HD2*)	3.3	1.5	1.0
	assign	(resid	29	and	name	HB2)	(resid	29	and	name	HA)	3.5	1.7	1.0
	assign	(resid	29	and	name	HB2)	(resid	29	and	name	HD1*)	3.6	1.8	0.8
	assign	(resid	29	and	name	HB2)	(resid	29	and	name	HN)	3.0	1.2	0.9
	assign	(resid	29	and	name	HD1*)	(resid	29	and	name	HG)	2.5	0.7	0.8
	assign	(resid	29	and	name	HD1*)	(resid	29	and	name	HN)	4.6	2.8	1.4
	assign	(resid	29	and	name	HD2*)	(resid	29	and	name	HN)	5.5	3.7	1.6
	assign	(resid	29	and	name	HG)	(resid	29	and	name	HD2*)	2.9	1.1	0.9
	assign	(resid	30	and	name	HA)	(resid	30	and	name	HB2)	2.9	1.1	0.9
	assign	(resid	30	and	name	HB2)	(resid	30	and	name	HG2)	2.7	0.9	0.8
150	assign	(resid	31	and	name	HA)	(resid	31	and	name	HB2)	3.3	1.5	1.0
	assign	(resid	31	and	name	HB1)	(resid	31	and	name	HA)	3.4	1.6	1.0
	assign	(resid	31	and	name	HB1)	(resid	31	and	name	HE2)	3.7	1.9	1.1
	assign	(resid	31	and	name	HB1)	(resid	31	and	name	HG2)	2.6	0.8	0.8
	assign	(resid	31	and	name	HB1)	(resid	31	and	name	HN)	3.3	1.5	1.0
	assign	(resid	31	and	name	HG2)	(resid	31	and	name	HA)	3.9	2.1	1.2
	assign	(resid	32	and	name	HA)	(resid	32	and	name	HN)	3.9	2.1	1.2
	assign	(resid	32	and	name	HB*)	(resid	32	and	name	HA)	3.4	1.6	1.0
	assign	(resid	32	and	name	HB*)	(resid	32	and	name	HN)	3.4	1.6	1.0
	assign	(resid	33	and	name	HA)	(resid	33	and	name	HD*)	3.6	1.8	1.1
160	assign	(resid	33	and	name	HA)	(resid	33	and	name	HN)	3.5	1.7	1.0
	assign	(resid	33	and	name	HB1)	(resid	33	and	name	HA)	3.7	1.9	1.1
	assign	(resid	33	and	name	HB1)	(resid	33	and	name	HD*)	2.8	1.0	0.8
	assign	(resid	33	and	name	HB1)	(resid	33	and	name	HE*)	4.2	2.4	1.3
	assign	(resid	33	and	name	HD*)	(resid	33	and	name	HN)	4.0	2.2	1.2
	assign	(resid	33	and	name	HN)	(resid	33	and	name	HB1)	3.5	1.7	1.1
	assign	(resid	34	and	name	HA)	(resid	34	and	name	HB1)	3.4	1.6	1.0
	assign	(resid	34	and	name	HA)	(resid	34	and	name	HB2)	3.3	1.5	1.0
	assign	(resid	34	and	name	HA)	(resid	34	and	name	HD1)	4.2	2.4	1.3
	assign	(resid	34	and	name	HA)	(resid	34	and	name	HN)	3.5	1.7	1.1
170	assign	(resid	34	and	name	HB1)	(resid	34	and	name	HE3)	5.1	3.3	1.5
	assign	(resid	34	and	name	HB1)	(resid	34	and	name	HZ3)	4.4	2.6	1.0
	assign	(resid	34	and	name	HB2)	(resid	34	and	name	HN)	3.5	1.7	1.0
	assign	(resid	34	and	name	HD1)	(resid	34	and	name	HB1)	4.1	2.3	1.2
	assign	(resid	34	and	name	HD1)	(resid	34	and	name	HB2)	4.0	2.2	1.2
	assign	(resid	34	and	name	HN)	(resid	34	and	name	HB1)	3.8	2.0	1.1
	assign	(resid	34	and	name	HN)	(resid	34	and	name	HD1)	4.2	2.4	1.3
	assign	(resid	34	and	name	HN)	(resid	34	and	name	HE3)	4.7	2.9	1.4
	assign	(resid	35	and	name	HA)	(resid	35	and	name	HB2)	3.6	1.8	1.1
	assign	(resid	35	and	name	HA)	(resid	35	and	name	HD1*)	3.5	1.7	1.1
180	assign	(resid	35	and	name	HA)	(resid	35	and	name	HN)	3.6	1.8	1.1
	assign	(resid	35	and	name	HB1)	(resid	35	and	name	HA)	3.5	1.7	1.1
	assign	(resid	35	and	name	HB1)	(resid	35	and	name	HG)	3.6	1.8	1.1
	assign	(resid	35	and	name	HB1)	(resid	35	and	name	HN)	3.8	2.0	1.1
	assign	(resid	35	and	name	HB2)	(resid	35	and	name	HB1)	2.9	1.1	0.9
	assign	(resid	35	and	name	HB2)	(resid	35	and	name	HN)	3.1	1.3	0.9
	assign	(resid	35	and	name	HD1*)	(resid	35	and	name	HB1)	3.1	1.3	0.9
	assign	(resid	35	and	name	HD1*)	(resid	35	and	name	HB2)	3.4	1.6	1.0
	assign	(resid	35	and	name	HD1*)	(resid	35	and	name	HN)	4.7	2.9	1.4
	assign	(resid	36	and	name	HA)	(resid	36	and	name	HB1)	4.0	2.2	1.2
190	assign	(resid	36	and	name	HA)	(resid	36	and	name	HB2)	4.4	2.6	1.3
	assign	(resid	36	and	name	HA)	(resid	36	and	name	HD*)	4.6	2.8	1.4
	assign	(resid	36	and	name	HB1)	(resid	36	and	name	HB2)	2.5	0.7	0.8
	assign	(resid	36	and	name	HB1)	(resid	36	and	name	HD*)	4.3	2.5	1.3
	assign	(resid	36	and	name	HB1)	(resid	36	and	name	HN)	3.7	1.9	1.1
	assign	(resid	36	and	name	HB2)	(resid	36	and	name	HD*)	4.4	2.6	1.3

## B. Ergebnisse

---

	assign	(resid	36	and	name	HB2)	(resid	36	and	name	HN)	4.0	2.2	1.2
	assign	(resid	36	and	name	HD*)	(resid	36	and	name	HN)	4.2	2.4	1.3
	assign	(resid	36	and	name	HN)	(resid	36	and	name	HA)	5.3	3.5	1.6
	assign	(resid	36	and	name	HN)	(resid	36	and	name	HE*)	4.7	2.9	1.4
200	sequentiell													
	assign	(resid	2	and	name	HA)	(resid	3	and	name	HN)	3.0	1.2	1.0
	assign	(resid	3	and	name	HA)	(resid	4	and	name	HN)	2.8	1.0	1.1
	assign	(resid	3	and	name	HB*)	(resid	4	and	name	HN)	4.3	2.5	1.3
	assign	(resid	4	and	name	HB*)	(resid	5	and	name	HN)	3.7	1.9	1.1
	assign	(resid	4	and	name	HA)	(resid	5	and	name	HN)	3.5	1.7	1.1
	assign	(resid	4	and	name	HN)	(resid	5	and	name	HN)	3.5	1.7	1.2
	assign	(resid	5	and	name	HA)	(resid	6	and	name	HN)	3.5	1.7	0.9
	assign	(resid	5	and	name	HB1)	(resid	6	and	name	HN)	4.1	2.3	1.2
210	assign	(resid	5	and	name	HB2)	(resid	6	and	name	HN)	4.1	2.3	1.2
	assign	(resid	5	and	name	HG2)	(resid	6	and	name	HN)	5.1	3.3	1.5
	assign	(resid	6	and	name	HB*)	(resid	7	and	name	HN)	3.0	1.2	1.1
	assign	(resid	6	and	name	HA)	(resid	7	and	name	HN)	3.6	1.8	1.2
	assign	(resid	7	and	name	HA)	(resid	8	and	name	HN)	2.8	1.0	0.9
	assign	(resid	7	and	name	HD1)	(resid	8	and	name	HN)	5.0	3.2	1.0
	assign	(resid	7	and	name	HB1)	(resid	8	and	name	HN)	3.2	1.4	1.0
	assign	(resid	7	and	name	HB2)	(resid	8	and	name	HN)	3.5	1.7	1.0
	assign	(resid	7	and	name	HD1)	(resid	8	and	name	HN)	4.8	3.0	1.0
	assign	(resid	7	and	name	HN)	(resid	8	and	name	HG*)	4.8	1.7	1.0
220	assign	(resid	8	and	name	HA)	(resid	9	and	name	HN)	3.0	1.2	0.8
	assign	(resid	8	and	name	HN)	(resid	9	and	name	HN)	3.8	2.0	1.0
	assign	(resid	8	and	name	HB2)	(resid	9	and	name	HN)	3.8	2.0	1.1
	assign	(resid	9	and	name	HB*)	(resid	10	and	name	HN)	3.4	1.6	1.0
	assign	(resid	9	and	name	HA)	(resid	10	and	name	HN)	2.5	0.7	0.9
	assign	(resid	9	and	name	HB*)	(resid	10	and	name	HA)	5.0	3.2	1.8
	assign	(resid	10	and	name	HB)	(resid	11	and	name	HN)	3.7	1.9	1.1
	assign	(resid	10	and	name	HG1*)	(resid	11	and	name	HN)	3.5	1.7	1.0
	assign	(resid	10	and	name	HA)	(resid	11	and	name	HN)	2.5	0.7	0.9
	assign	(resid	10	and	name	HN)	(resid	11	and	name	HN)	3.0	1.2	1.0
230	assign	(resid	12	and	name	HA*)	(resid	13	and	name	HN)	3.0	1.2	1.0
	assign	(resid	13	and	name	HA)	(resid	14	and	name	HN)	3.0	1.2	0.9
	assign	(resid	14	and	name	HA)	(resid	15	and	name	HN)	3.0	1.2	1.0
	assign	(resid	15	and	name	HA)	(resid	16	and	name	HN)	3.2	1.4	1.1
	assign	(resid	16	and	name	HN)	(resid	17	and	name	HN)	3.3	1.5	1.1
	assign	(resid	19	and	name	HA)	(resid	20	and	name	HN)	3.1	1.3	0.9
	assign	(resid	19	and	name	HN)	(resid	20	and	name	HN)	4.0	2.2	1.2
	assign	(resid	19	and	name	HB*)	(resid	20	and	name	HN)	4.0	2.2	0.9
	assign	(resid	20	and	name	HB1)	(resid	21	and	name	HN)	3.8	2.0	1.3
	assign	(resid	20	and	name	HB2)	(resid	21	and	name	HN)	3.8	2.0	1.3
240	assign	(resid	20	and	name	HA)	(resid	21	and	name	HN)	3.2	1.4	0.9
	assign	(resid	21	and	name	HN)	(resid	22	and	name	HD1)	4.2	2.9	1.4
	assign	(resid	21	and	name	HN)	(resid	22	and	name	HD2)	4.7	2.4	1.3
	assign	(resid	21	and	name	HB1)	(resid	22	and	name	HD1)	4.0	2.4	1.2
	assign	(resid	21	and	name	HB1)	(resid	22	and	name	HD2)	3.6	1.8	1.2
	assign	(resid	21	and	name	HD2*)	(resid	22	and	name	HD1)	4.5	2.7	1.4
	assign	(resid	21	and	name	HD2*)	(resid	22	and	name	HD2)	4.0	2.2	1.2
	assign	(resid	21	and	name	HB2)	(resid	22	and	name	HD1)	2.2	0.4	1.0
	assign	(resid	21	and	name	HB2)	(resid	22	and	name	HD2)	3.0	1.2	1.2
	assign	(resid	21	and	name	HA)	(resid	22	and	name	HG*)	3.9	2.1	1.2
250	assign	(resid	21	and	name	HA)	(resid	22	and	name	HD1)	2.6	0.8	0.9
	assign	(resid	21	and	name	HA)	(resid	22	and	name	HD2)	2.6	0.8	0.9
	assign	(resid	22	and	name	HA)	(resid	23	and	name	HB*)	4.0	2.2	1.0
	assign	(resid	22	and	name	HB2)	(resid	23	and	name	HN)	4.2	2.4	1.3
	assign	(resid	22	and	name	HD2)	(resid	23	and	name	HN)	4.9	3.1	1.5
	assign	(resid	22	and	name	HA)	(resid	23	and	name	HN)	2.2	0.4	0.8
	assign	(resid	23	and	name	HN)	(resid	24	and	name	HN)	3.5	1.7	1.1
	assign	(resid	23	and	name	HA)	(resid	24	and	name	HN)	3.5	1.7	0.9
	assign	(resid	23	and	name	HB*)	(resid	24	and	name	HN)	3.4	1.6	1.2
	assign	(resid	23	and	name	HD2*)	(resid	24	and	name	HN)	4.0	2.2	1.3
260	assign	(resid	23	and	name	HG)	(resid	24	and	name	HN)	3.4	1.6	1.0
	assign	(resid	24	and	name	HG2)	(resid	25	and	name	HN)	5.0	3.2	1.3
	assign	(resid	24	and	name	HN)	(resid	25	and	name	HN)	3.0	1.2	1.2
	assign	(resid	25	and	name	HB1)	(resid	26	and	name	HN)	3.1	1.3	1.2

## B. Ergebnisse

270	assign	(resid	25	and	name	HB2)	(resid	26	and	name	HN)	3.1	1.3	1.2
	assign	(resid	25	and	name	HA)	(resid	26	and	name	HN)	3.5	1.7	0.9
	assign	(resid	26	and	name	HA)	(resid	27	and	name	HN)	3.0	1.2	0.9
	assign	(resid	26	and	name	HN)	(resid	27	and	name	HN)	4.5	2.7	1.0
	assign	(resid	27	and	name	HA)	(resid	28	and	name	HN)	3.7	1.9	1.0
	assign	(resid	28	and	name	HB1)	(resid	29	and	name	HN)	4.5	2.7	1.4
	assign	(resid	28	and	name	HB2)	(resid	29	and	name	HN)	4.8	3.0	1.4
	assign	(resid	29	and	name	HD2*)	(resid	30	and	name	HN)	5.0	3.2	1.0
	assign	(resid	30	and	name	HB2)	(resid	31	and	name	HN)	3.5	1.7	1.2
	assign	(resid	30	and	name	HA)	(resid	31	and	name	HN)	3.1	1.3	0.9
	assign	(resid	30	and	name	HN)	(resid	31	and	name	HN)	4.0	2.2	1.0
	assign	(resid	31	and	name	HA)	(resid	32	and	name	HN)	2.8	1.0	0.9
280	assign	(resid	32	and	name	HN)	(resid	33	and	name	HN)	4.5	2.7	1.3
	assign	(resid	32	and	name	HA)	(resid	33	and	name	HN)	2.8	1.0	0.9
	assign	(resid	32	and	name	HB*)	(resid	33	and	name	HN)	2.8	2.0	1.2
	assign	(resid	32	and	name	HB*)	(resid	33	and	name	HD1)	4.5	2.7	1.4
	assign	(resid	32	and	name	HA)	(resid	33	and	name	HD1)	5.2	3.4	1.5
	assign	(resid	33	and	name	HD2)	(resid	34	and	name	HA)	4.0	2.2	1.0
	assign	(resid	33	and	name	HD2)	(resid	34	and	name	HN)	5.0	3.2	1.2
	assign	(resid	33	and	name	HA)	(resid	34	and	name	HN)	2.8	1.0	1.0
	assign	(resid	33	and	name	HN)	(resid	34	and	name	HN)	3.0	1.2	1.1
	assign	(resid	33	and	name	HB*)	(resid	34	and	name	HN)	3.2	1.4	1.0
	assign	(resid	34	and	name	HN)	(resid	35	and	name	HN)	3.5	1.7	1.0
	assign	(resid	34	and	name	HA)	(resid	35	and	name	HN)	2.5	0.7	0.9
290	assign	(resid	34	and	name	HB1)	(resid	35	and	name	HN)	3.8	2.0	1.3
	assign	(resid	34	and	name	HB2)	(resid	35	and	name	HN)	3.8	2.0	1.3
	assign	(resid	34	and	name	HD1)	(resid	35	and	name	HG)	5.5	3.7	1.6
	assign	(resid	34	and	name	HD1)	(resid	35	and	name	HD1*)	3.8	2.0	1.2
	assign	(resid	34	and	name	HD1)	(resid	35	and	name	HG)	4.5	2.7	1.0
	assign	(resid	35	and	name	HB2)	(resid	36	and	name	HN)	4.0	2.2	1.0
	assign	(resid	35	and	name	HB1)	(resid	36	and	name	HN)	4.0	2.2	1.0
	assign	(resid	35	and	name	HA)	(resid	36	and	name	HN)	2.8	1.0	1.0
	assign	(resid	35	and	name	HN)	(resid	36	and	name	HN)	3.0	1.2	1.1
	medium													
	assign	(resid	3	and	name	HN)	(resid	7	and	name	HB1)	4.5	2.7	1.3
	assign	(resid	5	and	name	HN)	(resid	7	and	name	HN)	4.8	3.0	1.0
assign	(resid	6	and	name	HN)	(resid	8	and	name	HN)	3.7	1.9	1.1	
assign	(resid	7	and	name	HA)	(resid	9	and	name	HN)	4.2	2.4	1.2	
assign	(resid	14	and	name	HB)	(resid	17	and	name	HN)	3.8	2.0	1.0	
assign	(resid	14	and	name	HG1)	(resid	17	and	name	HB*)	3.0	1.2	0.9	
assign	(resid	15	and	name	HN)	(resid	17	and	name	HB*)	4.8	3.0	1.5	
assign	(resid	16	and	name	HA)	(resid	19	and	name	HB*)	4.3	2.5	1.3	
assign	(resid	18	and	name	HG2)	(resid	21	and	name	HD1*)	4.6	2.8	1.0	
assign	(resid	22	and	name	HA)	(resid	24	and	name	HN)	4.8	3.0	1.4	
assign	(resid	22	and	name	HA)	(resid	25	and	name	HN)	4.8	3.0	1.0	
assign	(resid	22	and	name	HD2)	(resid	25	and	name	HB2)	4.0	2.2	1.0	
assign	(resid	22	and	name	HD1)	(resid	25	and	name	HB2)	5.0	3.2	1.0	
assign	(resid	23	and	name	HD1*)	(resid	27	and	name	HE2*)	4.0	2.2	1.0	
assign	(resid	23	and	name	HN)	(resid	25	and	name	HN)	3.6	1.8	1.2	
assign	(resid	23	and	name	HN)	(resid	26	and	name	HN)	4.8	3.0	1.3	
assign	(resid	25	and	name	HA)	(resid	29	and	name	HD2*)	4.7	2.9	1.0	
assign	(resid	27	and	name	HB2)	(resid	29	and	name	HN)	5.5	3.7	1.0	
assign	(resid	29	and	name	HB*)	(resid	33	and	name	HN)	4.3	2.5	1.3	
assign	(resid	32	and	name	HB*)	(resid	34	and	name	HN)	4.4	2.6	1.3	
assign	(resid	32	and	name	HA)	(resid	34	and	name	HN)	5.2	3.4	1.6	
assign	(resid	33	and	name	HA)	(resid	35	and	name	HN)	4.5	2.7	1.5	
assign	(resid	33	and	name	HN)	(resid	35	and	name	HN)	4.0	1.1	1.4	
assign	(resid	33	and	name	HD2)	(resid	35	and	name	HN)	4.5	2.7	1.3	
assign	(resid	34	and	name	HN)	(resid	36	and	name	HN)	4.0	2.2	1.0	
weitreichend														
assign	(resid	7	and	name	HD1)	(resid	30	and	name	HA)	4.8	3.0	1.3	
assign	(resid	7	and	name	HD1)	(resid	33	and	name	HD2)	5.9	4.1	1.8	
assign	(resid	7	and	name	HD2)	(resid	35	and	name	HB*)	3.5	2.1	1.9	
assign	(resid	7	and	name	HD2)	(resid	36	and	name	HN)	4.8	3.0	1.4	
assign	(resid	7	and	name	HE*)	(resid	33	and	name	HD2)	5.0	3.2	1.0	
assign	(resid	7	and	name	HE1)	(resid	34	and	name	HN)	4.0	2.2	1.1	

```

assign (resid 7 and name HZ) (resid 33 and name HD2) 5.0 3.2 1.5
assign (resid 7 and name HZ) (resid 36 and name HD1) 4.0 2.2 0.8
assign (resid 7 and name HE2) (resid 35 and name HA) 5.0 3.2 1.0
assign (resid 7 and name HD1) (resid 34 and name HN) 4.8 3.0 1.0
assign (resid 7 and name HZ) (resid 36 and name HN) 4.8 3.0 1.0
assign (resid 7 and name HD*) (resid 34 and name HN) 5.0 3.2 1.0
assign (resid 7 and name HD*) (resid 35 and name HN) 5.5 3.7 1.0
assign (resid 7 and name HE*) (resid 35 and name HN) 4.5 2.7 1.0
340 assign (resid 7 and name HE*) (resid 36 and name HB2) 5.0 3.2 1.0
assign (resid 21 and name HD1*) (resid 26 and name HA) 4.0 2.2 1.3
assign (resid 21 and name HD1*) (resid 29 and name HB2) 4.0 0.5 0.7
assign (resid 21 and name HB2) (resid 26 and name HE2*) 5.0 3.2 1.2
assign (resid 21 and name HB*) (resid 26 and name HN) 5.5 3.7 1.0

```

## TALOS\_constraints.tbl

```

! 1. Y 2 Phi -79.00 40.00 (-119.00 to -39.00)

assign (resid 1 and name C) (resid 2 and name N)
      (resid 2 and name CA) (resid 2 and name C) 1.0 -79.00 40.00 2

10 ! 2. A 4 Phi -61.00 14.00 (-75.00 to -47.00)

assign (resid 3 and name C) (resid 4 and name N)
      (resid 4 and name CA) (resid 4 and name C) 1.0 -61.00 14.00 2

! 3. E 5 Phi -63.49 13.38 (-76.87 to -50.11)

assign (resid 4 and name C) (resid 5 and name N)
      (resid 5 and name CA) (resid 5 and name C) 1.0 -63.49 13.38 2

20 ! 4. D 6 Phi -72.28 28.80 (-101.08 to -43.48)

assign (resid 5 and name C) (resid 6 and name N)
      (resid 6 and name CA) (resid 6 and name C) 1.0 -72.28 28.80 2

! 5. F 7 Phi -65.01 12.76 (-77.77 to -52.25)

assign (resid 6 and name C) (resid 7 and name N)
      (resid 7 and name CA) (resid 7 and name C) 1.0 -65.01 12.76 2

30 ! 6. K 8 Phi -64.00 14.00 (-78.00 to -50.00)

assign (resid 7 and name C) (resid 8 and name N)
      (resid 8 and name CA) (resid 8 and name C) 1.0 -64.00 14.00 2

! 7. G 12 Phi 93.00 28.00 (65.00 to 121.00)

assign (resid 11 and name C) (resid 12 and name N)
      (resid 12 and name CA) (resid 12 and name C) 1.0 93.00 28.00 2

40 ! 8. R 15 Phi -75.00 42.00 (-117.00 to -33.00)

assign (resid 14 and name C) (resid 15 and name N)
      (resid 15 and name CA) (resid 15 and name C) 1.0 -75.00 42.00 2

! 9. S 16 Phi -70.05 28.58 (-98.63 to -41.47)

assign (resid 15 and name C) (resid 16 and name N)
      (resid 16 and name CA) (resid 16 and name C) 1.0 -70.05 28.58 2

50 ! 10. A 17 Phi -80.00 34.00 (-114.00 to -46.00)

assign (resid 16 and name C) (resid 17 and name N)
      (resid 17 and name CA) (resid 17 and name C) 1.0 -80.00 34.00 2

! 11. K 18 Phi -64.00 8.00 (-72.00 to -56.00)

```

## B. Ergebnisse

---

```

assign (resid 17 and name C) (resid 18 and name N)
      (resid 18 and name CA) (resid 18 and name C) 1.0 -64.00 8.00 2

60 ! 12. A 19 Phi -73.07 38.16 (-111.23 to -34.91)

assign (resid 18 and name C) (resid 19 and name N)
      (resid 19 and name CA) (resid 19 and name C) 1.0 -73.07 38.16 2

! 13. N 20 Phi -81.00 64.00 (-145.00 to -17.00)

assign (resid 19 and name C) (resid 20 and name N)
      (resid 20 and name CA) (resid 20 and name C) 1.0 -81.00 64.00 2

70 ! 14. L 21 Phi -86.00 44.00 (-130.00 to -42.00)

assign (resid 20 and name C) (resid 21 and name N)
      (resid 21 and name CA) (resid 21 and name C) 1.0 -86.00 44.00 2

! 15. P 22 Phi -60.00 16.00 (-76.00 to -44.00)

!assign (resid 21 and name C) (resid 22 and name N)
      !(resid 22 and name CA) (resid 22 and name C) 1.0 -60.00 16.00 2

80 ! 16. L 23 Phi -60.00 14.00 (-74.00 to -46.00)

!assign (resid 22 and name C) (resid 23 and name N)
      !(resid 23 and name CA) (resid 23 and name C) 1.0 -60.00 14.00 2

! 17. E 24 Phi -66.30 10.26 (-76.56 to -56.04)

assign (resid 23 and name C) (resid 24 and name N)
      (resid 24 and name CA) (resid 24 and name C) 1.0 -66.30 10.26 2

90 ! 18. K 25 Phi -71.04 24.24 (-95.28 to -46.80)

assign (resid 24 and name C) (resid 25 and name N)
      (resid 25 and name CA) (resid 25 and name C) 1.0 -71.04 24.24 2

! 19. Q 26 Phi -64.01 9.34 (-73.35 to -54.67)

assign (resid 25 and name C) (resid 26 and name N)
      (resid 26 and name CA) (resid 26 and name C) 1.0 -64.01 9.34 2

100 ! 20. Q 27 Phi -63.70 9.28 (-72.98 to -54.42)

assign (resid 26 and name C) (resid 27 and name N)
      (resid 27 and name CA) (resid 27 and name C) 1.0 -63.70 9.28 2

! 21. N 28 Phi -64.00 26.00 (-90.00 to -38.00)

assign (resid 27 and name C) (resid 28 and name N)
      (resid 28 and name CA) (resid 28 and name C) 1.0 -64.00 26.00 2

110 ! 22. K 30 Phi -62.12 18.86 (-80.98 to -43.26)

assign (resid 29 and name C) (resid 30 and name N)
      (resid 30 and name CA) (resid 30 and name C) 1.0 -62.12 18.86 2

! 23. K 31 Phi -63.00 12.00 (-75.00 to -51.00)

assign (resid 30 and name C) (resid 31 and name N)
      (resid 31 and name CA) (resid 31 and name C) 1.0 -63.00 12.00 2

120 ! 24. A 32 Phi -62.00 8.00 (-70.00 to -54.00)

assign (resid 31 and name C) (resid 32 and name N)
      (resid 32 and name CA) (resid 32 and name C) 1.0 -62.00 8.00 2

```

## B. Ergebnisse

---

```

! 25. Y 33 Phi -79.63 40.94 (-120.57 to -38.69)

assign (resid 32 and name C) (resid 33 and name N)
      (resid 33 and name CA) (resid 33 and name C) 1.0 -79.63 40.94 2

130 ! 26. W 34 Phi -83.00 26.00 (-109.00 to -57.00)

assign (resid 33 and name C) (resid 34 and name N)
      (resid 34 and name CA) (resid 34 and name C) 1.0 -83.00 26.00 2

! 27. L 35 Phi -69.00 28.00 (-97.00 to -41.00)

assign (resid 34 and name C) (resid 35 and name N)
      (resid 35 and name CA) (resid 35 and name C) 1.0 -69.00 28.00 2

140
! 28. Y 2 Psi -24.00 54.00 (-78.00 to 30.00)

assign (resid 2 and name N) (resid 2 and name CA)
      (resid 2 and name C) (resid 3 and name N) 1.0 -24.00 54.00 2

! 29. A 4 Psi -43.00 12.00 (-55.00 to -31.00)

assign (resid 4 and name N) (resid 4 and name CA)
      (resid 4 and name C) (resid 5 and name N) 1.0 -43.00 12.00 2

150
! 30. E 5 Psi -39.97 8.78 (-48.75 to -31.19)

assign (resid 5 and name N) (resid 5 and name CA)
      (resid 5 and name C) (resid 6 and name N) 1.0 -39.97 8.78 2

! 31. D 6 Psi -33.61 27.38 (-60.99 to -6.23)

assign (resid 6 and name N) (resid 6 and name CA)
      (resid 6 and name C) (resid 7 and name N) 1.0 -33.61 27.38 2

160
! 32. F 7 Psi -34.91 34.00 (-68.91 to -0.91)

assign (resid 7 and name N) (resid 7 and name CA)
      (resid 7 and name C) (resid 8 and name N) 1.0 -34.91 34.00 2

! 33. K 8 Psi -35.00 30.00 (-65.00 to -5.00)

assign (resid 8 and name N) (resid 8 and name CA)
      (resid 8 and name C) (resid 9 and name N) 1.0 -35.00 30.00 2

170
! 34. G 12 Psi -8.00 30.00 (-38.00 to 22.00)

assign (resid 12 and name N) (resid 12 and name CA)
      (resid 12 and name C) (resid 13 and name N) 1.0 -8.00 30.00 2

! 35. R 15 Psi -24.00 48.00 (-72.00 to 24.00)

assign (resid 15 and name N) (resid 15 and name CA)
      (resid 15 and name C) (resid 16 and name N) 1.0 -24.00 48.00 2

180
! 36. S 16 Psi -30.65 25.52 (-56.17 to -5.13)

assign (resid 16 and name N) (resid 16 and name CA)
      (resid 16 and name C) (resid 17 and name N) 1.0 -30.65 25.52 2

! 37. A 17 Psi -20.00 46.00 (-66.00 to 26.00)

assign (resid 17 and name N) (resid 17 and name CA)
      (resid 17 and name C) (resid 18 and name N) 1.0 -20.00 46.00 2

190
! 38. K 18 Psi -39.00 12.00 (-51.00 to -27.00)

```



## B. Ergebnisse

---

```

assign (resid 18 and name N) (resid 18 and name CA)
      (resid 18 and name C) (resid 19 and name N) 1.0 -39.00 12.00 2

! 39. A 19 Psi -20.62 50.92 (-71.54 to 30.30)

assign (resid 19 and name N) (resid 19 and name CA)
      (resid 19 and name C) (resid 20 and name N) 1.0 -20.62 50.92 2
200 ! 40. N 20 Psi -29.00 34.00 (-63.00 to 5.00)

assign (resid 20 and name N) (resid 20 and name CA)
      (resid 20 and name C) (resid 21 and name N) 1.0 -29.00 34.00 2

! 41. L 21 Psi 121.00 62.00 (59.00 to 183.00)

!assign (resid 21 and name N) (resid 21 and name CA)
!      (resid 21 and name C) (resid 22 and name N) 1.0 121.00 62.00 2
210 ! 42. P 22 Psi 145.00 26.00 (119.00 to 171.00)

!assign (resid 22 and name N) (resid 22 and name CA)
!      (resid 22 and name C) (resid 23 and name N) 1.0 145.00 26.00 2

! 43. L 23 Psi -32.00 38.00 (-70.00 to 6.00)

assign (resid 23 and name N) (resid 23 and name CA)
      (resid 23 and name C) (resid 24 and name N) 1.0 -32.00 38.00 2
220 ! 44. E 24 Psi -42.07 13.54 (-55.61 to -28.53)

assign (resid 24 and name N) (resid 24 and name CA)
      (resid 24 and name C) (resid 25 and name N) 1.0 -42.07 13.54 2

! 45. K 25 Psi -29.87 33.64 (-63.51 to 3.77)

assign (resid 25 and name N) (resid 25 and name CA)
      (resid 25 and name C) (resid 26 and name N) 1.0 -29.87 33.64 2
230 ! 46. Q 26 Psi -42.77 27.36 (-70.13 to -15.41)

assign (resid 26 and name N) (resid 26 and name CA)
      (resid 26 and name C) (resid 27 and name N) 1.0 -42.77 27.36 2

! 47. Q 27 Psi -39.87 20.40 (-60.27 to -19.47)

assign (resid 27 and name N) (resid 27 and name CA)
      (resid 27 and name C) (resid 28 and name N) 1.0 -39.87 20.40 2
240 ! 48. N 28 Psi -41.00 12.00 (-53.00 to -29.00)

assign (resid 28 and name N) (resid 28 and name CA)
      (resid 28 and name C) (resid 29 and name N) 1.0 -41.00 12.00 2

! 49. K 30 Psi -38.15 34.06 (-72.21 to -4.09)

assign (resid 30 and name N) (resid 30 and name CA)
      (resid 30 and name C) (resid 31 and name N) 1.0 -38.15 34.06 2
250 ! 50. K 31 Psi -42.00 10.00 (-52.00 to -32.00)

assign (resid 31 and name N) (resid 31 and name CA)
      (resid 31 and name C) (resid 32 and name N) 1.0 -42.00 10.00 2

! 51. A 32 Psi -42.00 10.00 (-52.00 to -32.00)

assign (resid 32 and name N) (resid 32 and name CA)
      (resid 32 and name C) (resid 33 and name N) 1.0 -42.00 10.00 2

```

260

! 52. Y 33 Psi -28.49 29.82 (-58.31 to 1.33)

```
assign (resid 33 and name N) (resid 33 and name CA)
      (resid 33 and name C) (resid 34 and name N) 1.0 -28.49 29.82 2
```

! 53. W 34 Psi -19.00 48.00 (-67.00 to 29.00)

```
assign (resid 34 and name N) (resid 34 and name CA)
      (resid 34 and name C) (resid 35 and name N) 1.0 -19.00 48.00 2
```

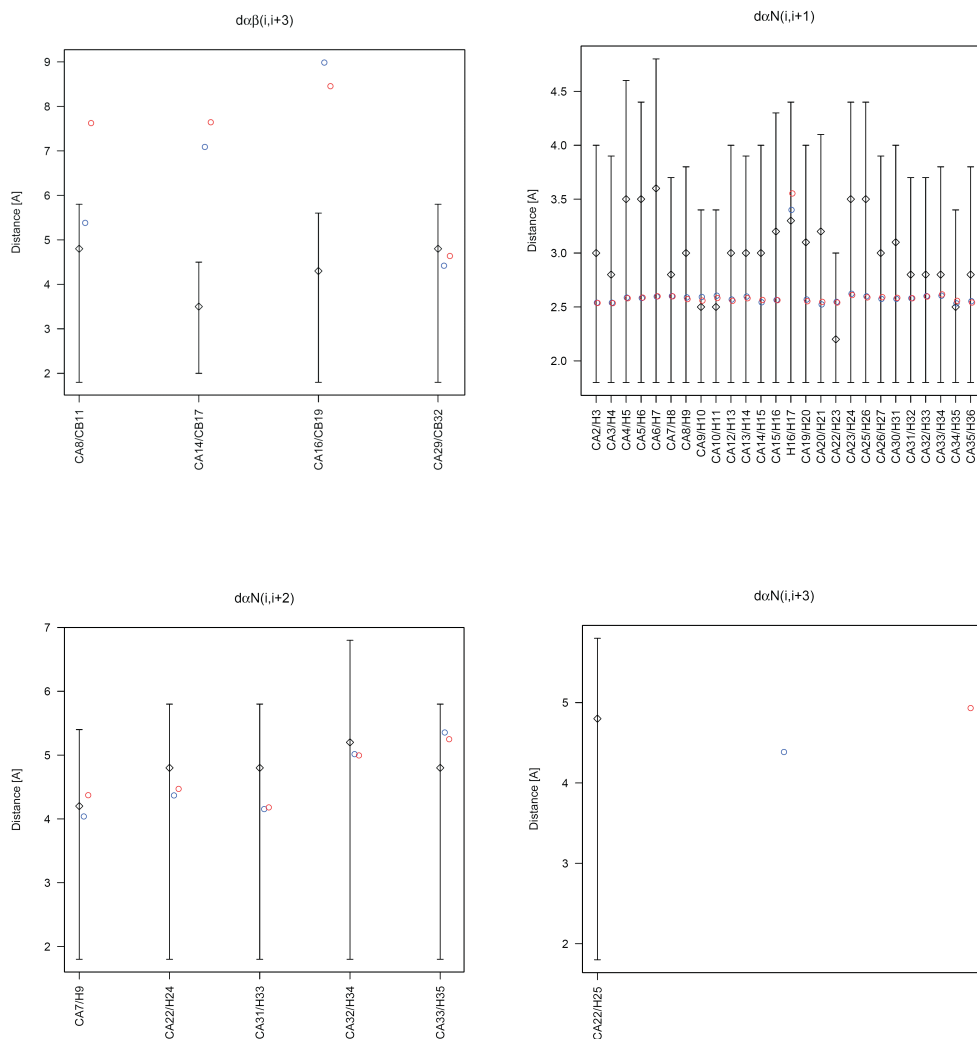
270

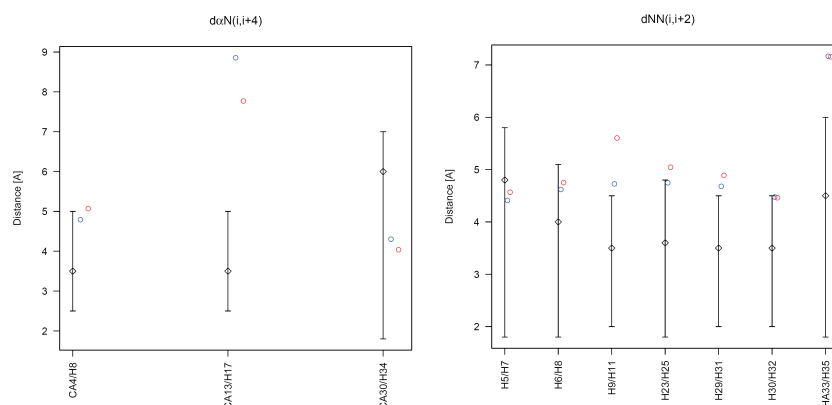
! 54. L 35 Psi -34.00 22.00 (-56.00 to -12.00)

```
assign (resid 35 and name N) (resid 35 and name CA)
      (resid 35 and name C) (resid 36 and name N) 1.0 -34.00 22.00 2
```

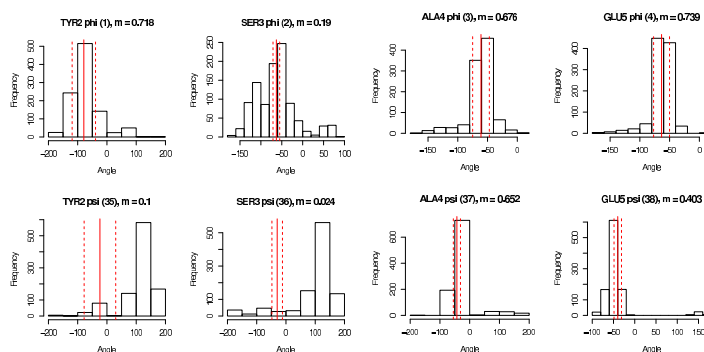
## Ergebnisse aus der Bioinformatik

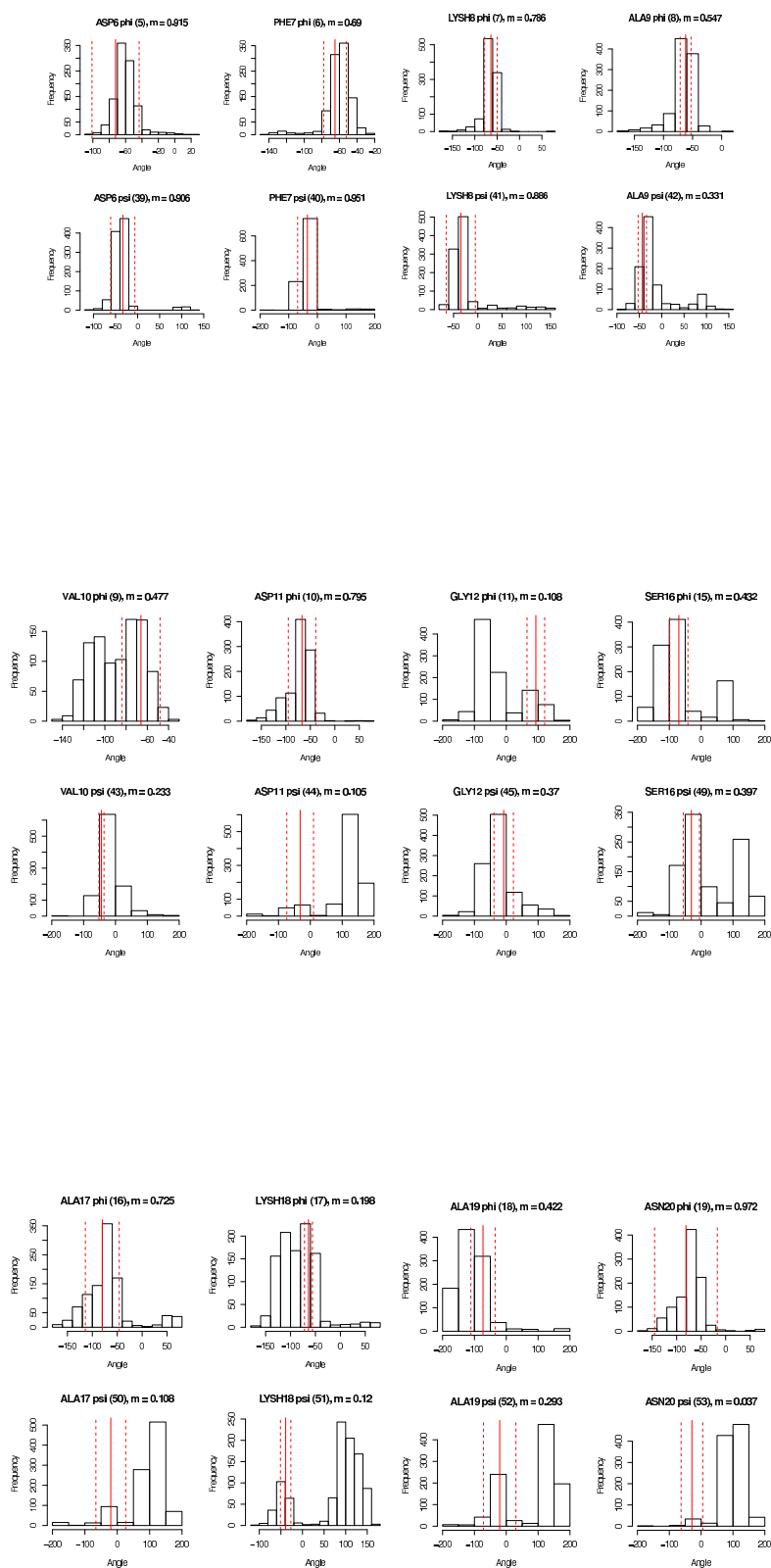
Berechnete Distanzen zwischen zwei Protonen für Trajektorien 1 und 2 (blau, rot) im Vergleich zu den über NMR ermittelten Daten.

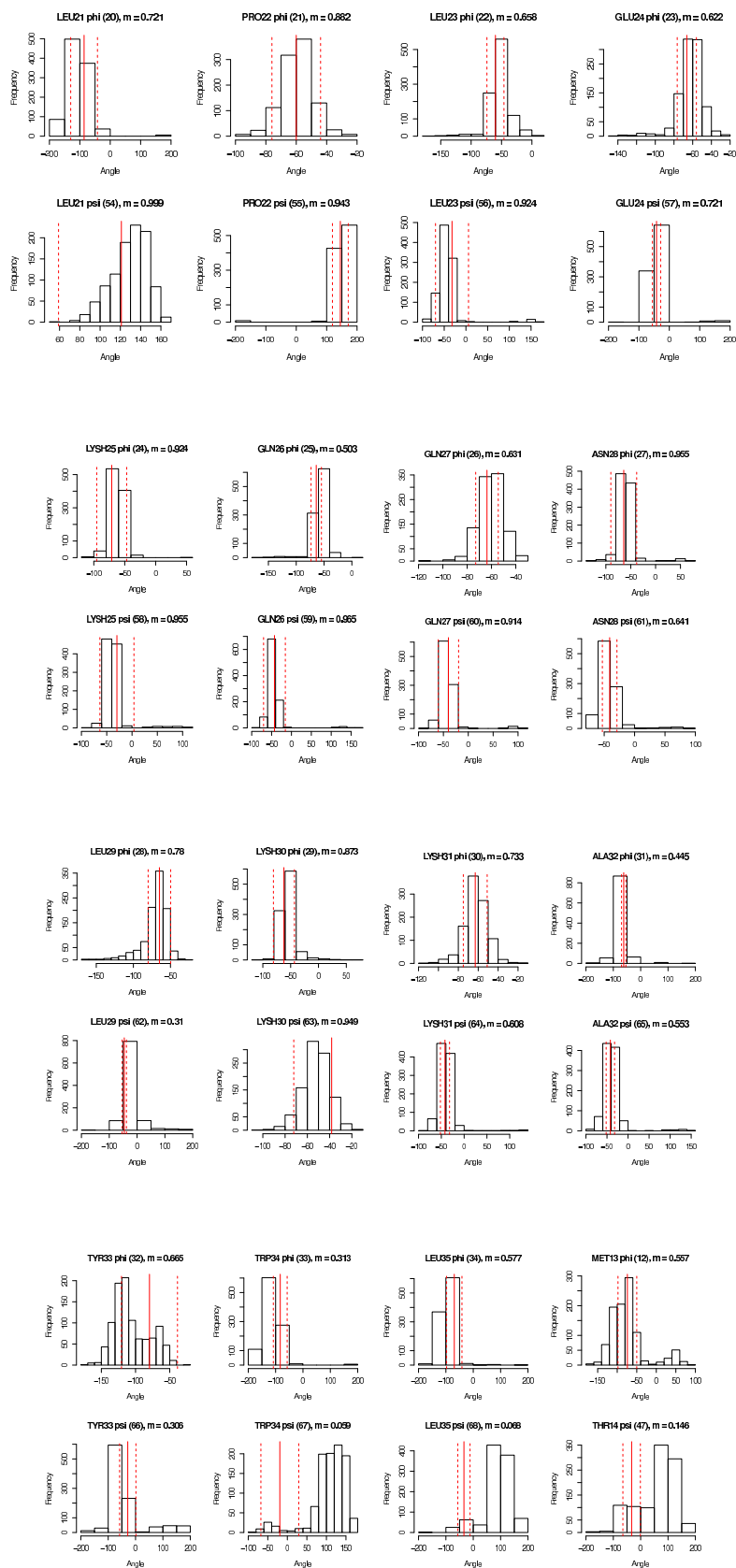


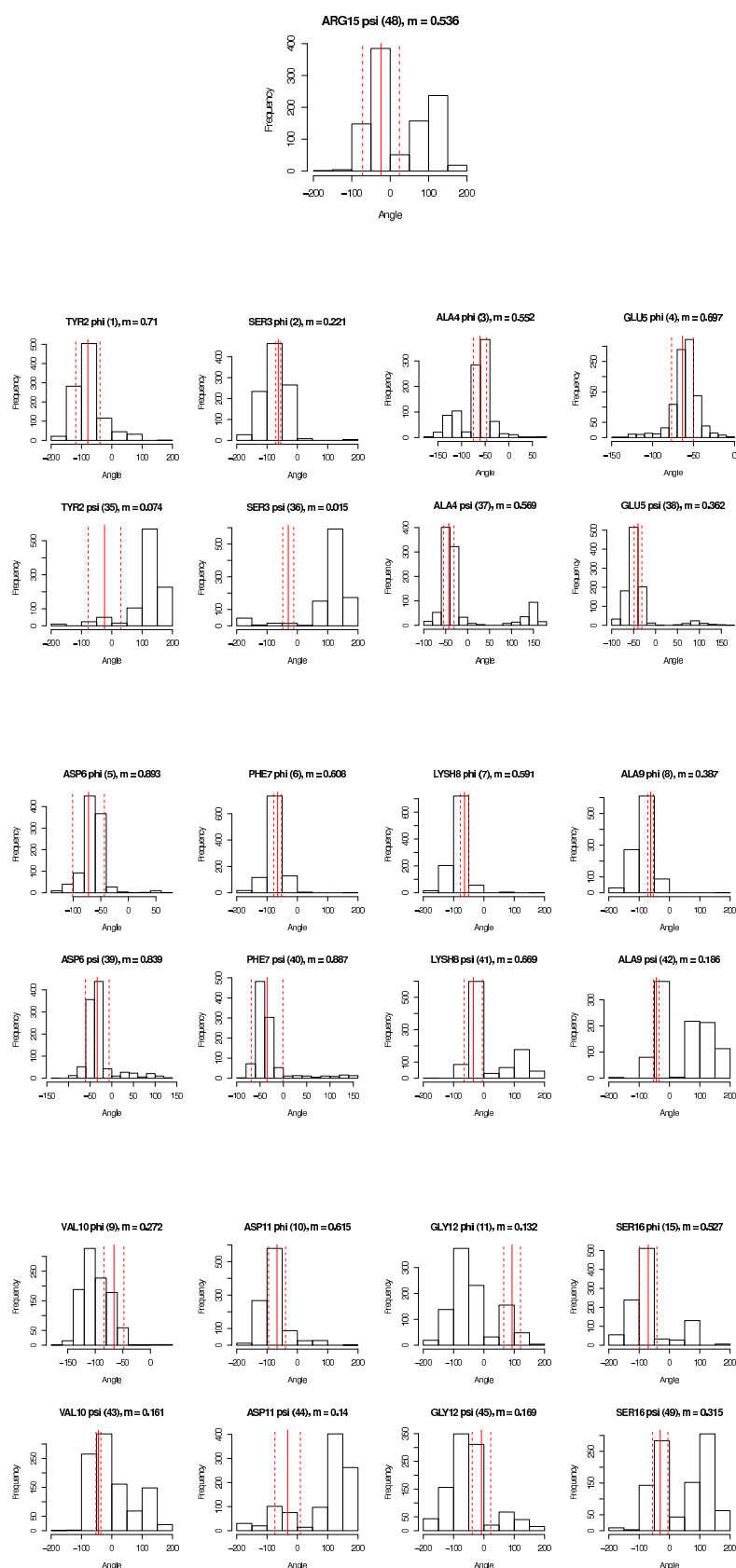


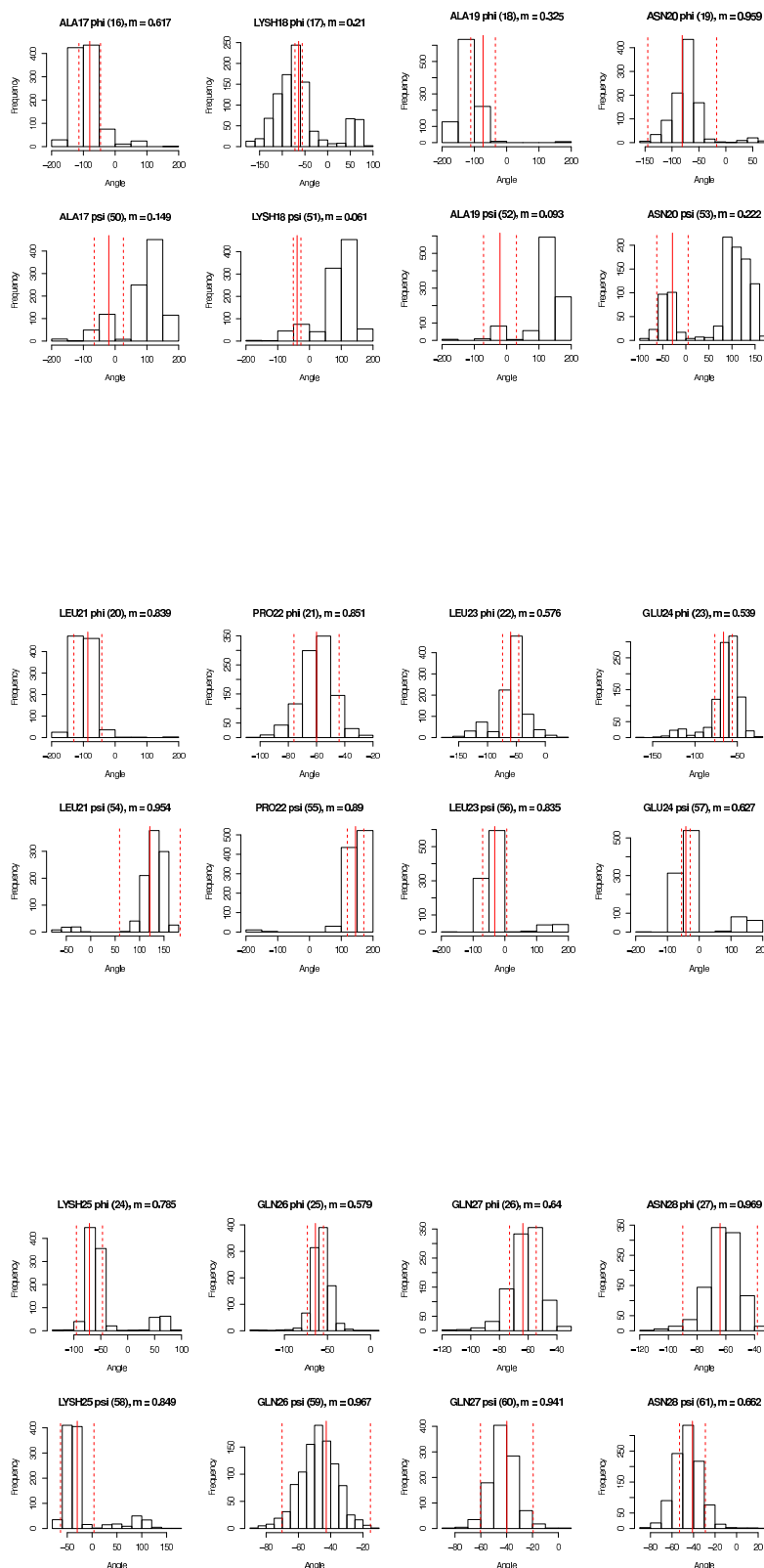
Statistische Diederwinkelverteilung für new2 aus Trajektorien 1 und 2 im Vergleich zu den für mut8 ermittelten Werten (rote, durchgezogene Linien). Die rot gestrichelten Linien stellen die erlaubte Abweichung der Werte für mut8 dar.  $m$  ist der Anteil an Strukturen, die sich in der jeweiligen Trajektorie innerhalb des von den beiden gestrichelten Linien eingegrenzten Bereiches befinden.

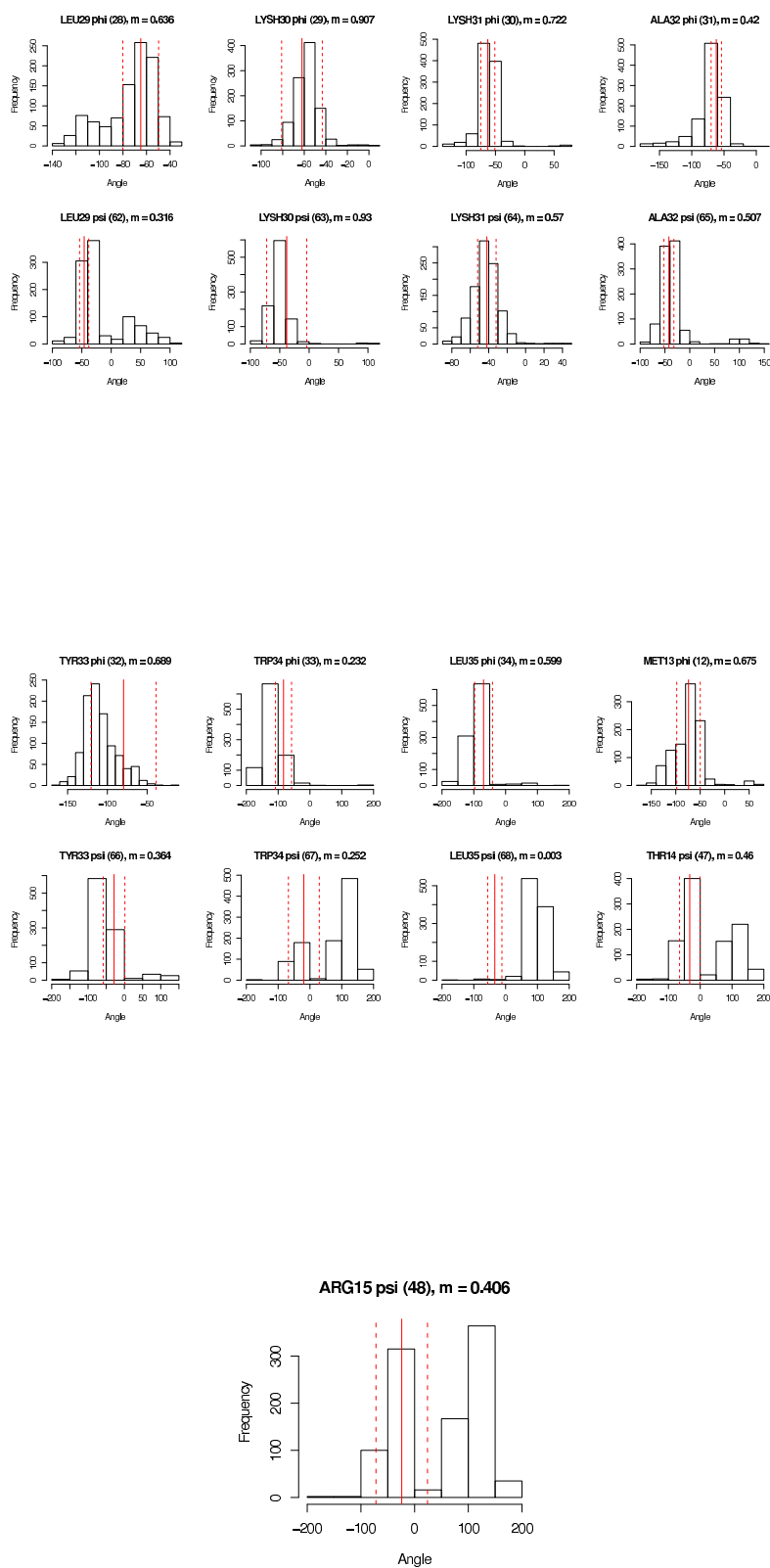














# Danksagung

Ich danke Prof. Dr. Peter Bayer für die Vergabe dieser Arbeit, für seine Geduld, sein Interesse und sein Vertrauen. Dank ihm habe ich diese Form der Strukturaufklärung von Proteinen und Peptiden für mich als *high-end*-Sudoku entdeckt.

Prof. Dr. Daniel Hoffmann gilt ebenfalls mein Dank für die Mit-Bereitstellung dieses Themas und für das freundliche Bemühen, einen Nicht-Bioinformatiker mit den faszinierenden evolutionären Algorithmen vertraut zu machen.

Ohne Manuel Prinz wäre ich wahrscheinlich verloren gewesen und wäre sowohl an Linux als auch an der Bioinformatik gescheitert. Ihm gilt ein ganz besonders großes Dankeschön für die ganze Hilfe und die hervorragende Kooperationsarbeit.

Der besten Arbeitsgruppe ein herzliches Dankeschön. Tina Stratmann für die Expertise beim Klonieren und das Bündnis gegen alle Aufräum- und Putzmuffel. Alma Rüppel für die Hilfe beim Expressieren und Reinigen. Peter Binz für die ganzen NMR-Tüfteleien. Dr. Jonathan Müller für die pffiffigen Ideen und die guten Gespräche. Akuma Saningong für den Kamerun4AfrikaClub, einmal Zukunft, bitte! Allen anderen danke ich für das feundschaftliche Miteinander.

Ricarda Maltaner, Nils Rohde und Robert Schwarz danke ich für die tolle Zusammenarbeit und die Möglichkeit, über mein eigenes Projekt hinaus noch andere Themengebiete zu behandeln.

Señor Dr. Sinnen danke ich ganz herzlich für die geduldige Hilfe am Computer und für die besten Sprüche im Spanischkurs, die diesen ungemein aufwerteten.

Dr. Mäl *in spe*, danke für die Annektierung Griechenlands, Schottlands und Frankreichs. Für die NMR-*sessions* und für die Musterschülerin im Spanischkurs für Profis. Prachtvoll.

Es gibt wohl keine andere Germanistik-Doktorandin außer Angela Kuhn, mit der man an einem Samstag Morgen NMR-Spektren auswerten kann und die darin genauso gut ist wie MedBio-Studenten im 7. Semester. Ihr ein ganz herzliches Dankeschön für das schnelle Korrekturlesen.

Der größte Dank geht an an meine Familie und an meinen Freund André. Für viele schöne Abende, die Profi-Koch Abendessen, das Abholen von der Uni (wenns mal wieder länger dauerte) und das nächtelange, geduldige Ausharren am NMR. Für deine und eure Liebe. Ganz besonders danken möchte ich meinem Bruder und Tex-Experten Manni, der mich davor bewahrte, kurzfristig doch in Word zu schreiben.



Wenn man alle logischen Lösungen eines Problems eliminiert, ist die Unlogische - obwohl unmöglich - unweigerlich richtig.

Das war mir entgangen, Holmes.

Das war mir klar, Watson.

# Lebenslauf

## Persönliche Daten

Name	Anja Matena
Geburtsort und -tag	Essen, 16.09.1983
Staatsangehörigkeit	deutsch

## Schulausbildung

1989 bis 1993	Grundschule St. Joseph in Essen-Horst
1993 bis 2002	Carl-Humann Gymnasium in Essen-Steele
	Abschluss: Abitur (Notendurchschnitt: 1,8)

## Hochschulstudium

10.2002	Immatrikulation an der Ruhr-Universität Bochum im Fach Biologie
09.2004	Vordiplom (Gesamtnote: 1,9)
10.2004	Praktikum am Klinikum der Universität Duisburg-Essen im Institut für Virologie
04.2006 - 01.2007	Diplomarbeit zum Thema „Rekonstitution und spektroskopische Untersuchung der Ligandenbindung von CopB aus <i>Sulfolobus solfataricus</i> “ (Note: 1,3) Abschluss: Diplombiologin (Gesamtnote: 1,2)
07.2008 - 03.2010	Teilnahme am Mentoring <sup>3</sup>
03.2007 - 07.2010	Doktorandin am Zentrum für Medizinische Biotechnologie (ZMB) der Universität Duisburg-Essen, AG Prof. Dr. Peter Bayer

# Erklärungen

Hiermit erkläre ich, gem. § 6 Abs. 2, Nr. 7 der Promotionsordnung der Math.-Nat.-Fachbereiche zur Erlangung der Dr. rer. nat., dass ich das Arbeitsgebiet, dem das Thema „Entwicklung und Strukturaufklärung funktionaler Protein-Domänen“ zuzuordnen ist, in Forschung und Lehre vertrete und den Antrag von Anja Matena befürworte.

Essen, den 22. April 2010

\_\_\_\_\_  
Unterschrift eines Mitglieds der Universität Duisburg-Essen

Hiermit erkläre ich, gem. § 6 Abs. 2, Nr. 6 der Promotionsordnung der Math.-Nat.-Fachbereiche zur Erlangung des Dr. rer. nat., dass ich die vorliegende Dissertation selbständig verfasst und mich keiner anderen als der angegebenen Hilfsmittel bedient habe.

Essen, den 22. April 2010

\_\_\_\_\_  
Unterschrift der Doktorandin

Hiermit erkläre ich, gem. § 6 Abs. 2, Nr. 8 der Promotionsordnung der Math.-Nat.-Fachbereiche zur Erlangung des Dr. rer. nat., dass ich keine anderen Promotionen bzw. Promotionsversuche in der Vergangenheit durchgeführt habe und dass diese Arbeit von keiner anderen Fakultät/Fachbereich abgelehnt worden ist.

Essen, den 22. April 2010

\_\_\_\_\_  
Unterschrift der Doktorandin